

DNA METHYLATION PROFILING OF FISH TUMOURS

by

LEDA MIRBAHAI

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Biosciences
The University of Birmingham
March 2012

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

Assessment of disease status in fish is used in the UK Clean Seas Environmental Monitoring Programme (CSEMP) as an indicator of the biological effects of contaminants in the marine environment. At some UK offshore sites the prevalence of liver tumours in *Limanda limanda* (dab) exceeds 20%. However, the molecular mechanisms of tumour formation and the causative agents are not known. The contribution of epigenetic mechanisms such as DNA methylation, although well-established in human tumourigenesis, is under-studied in tumours of aquatic species used in environmental monitoring or in modeling of human disease. In this thesis gene-specific (e.g. bisulfite sequencing PCR, real time-PCR) and high throughput techniques (e.g. methylated-DNA immunoprecipitation combined with microarrays or high throughput DNA sequencing, triple stage quadrupole tandem mass spectrometry) were used to investigate alterations in DNA methylation profiles of tumours of two fish species: the model species zebrafish (*Danio rerio*) with chemically induced liver tumours and the un-sequenced marine flatfish dab that develop liver tumours in the wild. The data presented provided a comprehensive characterisation of the DNA methylation pattern in zebrafish liver and the first evidence of alterations in the DNA methylation profiles of key genes in tumourigenesis pathways in any aquatic species. This suggested the possibility of the involvement of DNA methylation alterations in dab tumourigenesis. It was then demonstrated a statistically significant lower level (1.8 fold) of global DNA methylation in hepatocellular adenoma (HCA) and non-cancerous surrounding liver tissue (ST) compared to liver of non-cancer bearing dab. Components of the one-carbon cycle, the metabolic pathway involved in regulation of DNA methylation, were altered in both HCA and ST compared to healthy dab liver samples. In particular, *S*-adenosylhomocysteine levels, an inhibitor of DNA methylation reactions, were statistically significantly increased in HCA and ST compared to healthy dab liver samples. Based on the identified changes in the metabolites, methylome and transcriptome of ST, our study supports the epigenetic progenitor model of cancer. Furthermore, the evidence presented in these chapters suggests that chronic exposure to a mixture of pollutants contribute to global DNA hypomethylation followed by further epigenetic and genomic changes, especially in the non-genomic and genomic estrogenic pathways, leading to the development of tumours in dab. These findings suggest a link between the environment, epigenome and cancer in fish tumours in the wild.

Acknowledgments

I would like to thank my supervisors Professor Kevin Chipman and Dr Tim Williams for their guidance, support and motivating discussions throughout my PhD. In particular, I am grateful to Kevin who gave me the freedom to develop and follow my own ideas. Also, I would like to acknowledge NERC for funding my PhD and NBAF for funding the metabolomics studies. In particular, I am grateful to Dr Ulf Sommer and Professor Mark Viant for their expertise and assistance with optimising the LC-TSQ method used in Chapter 5. In addition, I would like to thank Dr Brett Lyons and John Bignell (Cefas, Weymouth) for providing the dab liver samples, Dr Li Ning and Guangliang Yin (Beijing Genomics Institute) for conducting the MeDIP-HTS, Professor Zhiyuan Gong (National University of Singapore) for providing the zebrafish liver samples and Lorraine Wallace (The Functional Genomics and Proteomic Unit, Birmingham) for her assistance during my PhD.

Also, I would like to thank Dr Anne Pheasant for her encouragement during my studies towards an MSc degree in Toxicology. Without her recommendation I would have never had the opportunity to work on this interesting project. Also, I would like to thank Anne for the card that she gave me after demonstrating for a practical session, which made me happier than anyone could imagine!

My work and my time on the 4th floor would have not been as enjoyable as it has been without the enthusiasm and friendly environment created by the members of the 4th floor and in specific Prof Kevin Chipman and Dr Nik Hodges groups. Especially, I would like to thank Dr Bob Harris for providing me with endless supplies of carrot cake and biscuits throughout my PhD. Also, I am grateful to Annette Evans for organising meetings with Kevin.

Many thanks to Rhiannon David, Huw Jones, Rachael Kershaw, Louise Stone, Nadine Taylor and Chibuzor Uchea for the non-science/non-football themed “tea breaks” and “outings”, which made my PhD most enjoyable.

Last but not least, I would like to thank my amazing and most wonderful parents, Sandra and Farrokh, for always supporting me. Their interest in my work and their encouragement has always inspired me. I could never thank them enough. Finally, I would like to thank my partner in crime, my idol, my sister, Ladan, for introducing me to the fascinating world of science and showing me how to set up experiments with powder, mud and glue when we were

kids, for never getting bored of me following her footsteps, for listening to my complains after a bad science day, and for being the kindest person I have ever met!

List of contents

Chapter 1: General introduction	1
1.1. General introduction.....	2
1.2. Epigenetics.....	2
1.3. DNA methylation.....	4
1.3.1. DNA methyltransferases; <i>de novo</i> and maintenance DNA methylation.....	5
1.3.2. DNA methylation and regulation of gene expression: Crosstalk between DNA methylation and histone modifications.....	13
1.3.2.1. DNA methylation as a gene silencing mechanism.....	13
1.3.2.2. DNA methylation in relation to active transcription.....	18
1.3.2.3. General inverse correlation between gene expression and DNA methylation...	23
1.3.3. The one-carbon cycle.....	24
1.3.4. Epigenetic regulation of normal functions in the cells.....	27
1.3.4.1. Imprinting.....	27
1.3.4.2. X-chromosome inactivation.....	28
1.3.4.3. Differentiation and tissue specific gene expression.....	29
1.3.4.4. Transposons.....	29
1.3.5. Cancer as an epigenetic disease.....	30
1.3.5.1. DNA methylation and cancer.....	30
1.3.5.1.1. DNA hypomethylation.....	32
1.3.5.1.2. DNA hypermethylation and inactivation of tumour suppressor genes.....	36
1.3.5.1.3. Increased rate of mutations.....	37
1.3.5.2. Epigenetic mechanisms and cancer: An interface between the environment and the genome.....	38
1.3.5.3. Two models of tumourigenesis: The classic multistep model initiated by mutation and the epigenetic progenitor model.....	46
1.4. DNA methylation and its implication in marine biology.....	51
1.4.1. Cancer in fish.....	52
1.4.1.1. Use of fish in carcinogenicity studies.....	52
1.4.1.2. Common dab (<i>Limanda limanda</i>) and environmental carcinogenicity studies..	53
1.4.1.3. Zebrafish (<i>Danio rerio</i>) as a laboratory cancer model.....	56
1.5. Aims.....	57
Chapter 2: Material and Methods	59
2.1. Chemicals.....	60
2.2. Test organisms.....	60
2.2.1. Zebrafish (<i>Danio rerio</i>).....	60
2.2.1.1. Chemical induction of hepatocellular carcinoma in zebrafish liver.....	60
2.2.1.2. DNA extraction from zebrafish liver samples.....	61
2.2.2. Dab (<i>Limanda limanda</i>).....	62
2.2.2.1. Collection of dab livers.....	62
2.2.2.2. Histopathology.....	66
2.2.2.3. Sample preparation.....	68

2.2.2.3.1. DNA and RNA extraction.....	69
2.2.2.3.2. Extraction of metabolites.....	70
2.2.3. European flounder (<i>Platichthys flesus</i>) liver tissue samples and Calf (<i>Bos taurus</i>) thymus DNA.....	71
2.3. Global measurement of DNA methylation levels.....	71
2.3.1. Enzyme-linked immunosorbent assay (ELISA)-based method.....	71
2.3.2. HPLC and LC-MS/MS.....	72
2.3.2.1. Removal of RNA contamination of DNA samples for both HPLC and LC MS/MS.....	72
2.3.2.2. DNA hydrolysis prior to HPLC.....	73
2.3.2.3. DNA hydrolysis prior to LC-MS/MS.....	73
2.3.2.4. Standards.....	74
2.3.2.5. Reversed-phase high-performance liquid chromatography.....	74
2.3.2.6. Nucleotide detection.....	75
2.3.2.7. LC-MS/MS.....	75
2.4. Methylated-DNA immunoprecipitation (MeDIP).....	76
2.5. Microarray experiments.....	78
2.5.1. Pilot study: cDNA flounder microarray.....	78
2.5.1.1. DNA labelling and purification.....	80
2.5.1.2. Hybridisation.....	81
2.5.1.3. Washing and scanning.....	81
2.5.2. CGI (1.5kb downstream to 1kb upstream of TSS) zebrafish tiling microarray experiment.....	82
2.5.2.1. Design of the 4x44k format CGI (1.5kb downstream to 1kb upstream of TSS) zebrafish tiling microarray.....	83
2.5.2.2. DNA labelling.....	83
2.5.2.3. Hybridisation.....	84
2.5.2.4. Washing and scanning.....	85
2.5.3. Dab MeDIP <i>de novo</i> high-throughput sequencing and transcriptomic profiling..	87
2.5.3.1. MeDIP <i>de novo</i> high-throughput sequencing.....	89
2.5.3.2. Identification and annotation of differentially methylated regions.....	90
2.5.3.3. Design of flatfish dab specific 8x15k oligo microarray based on contigs achieved from MeDIP <i>de novo</i> HTS.....	90
2.5.3.4. Quality assessment.....	91
2.5.3.5. DNase treatment of RNA samples.....	95
2.5.3.6. cDNA and cRNA synthesis and labelling.....	95
2.5.3.7. Hybridisation.....	96
2.5.3.8. Washing and scanning.....	97
2.6. Bisulfite sequencing PCR.....	97
2.6.1. Overview of the bisulfite sequencing PCR.....	97
2.6.2. Sodium bisulfite treatment.....	98
2.6.3. Design of the bisulfite sequencing PCR primers.....	98
2.6.4. Amplification of the sodium bisulfite treated DNA.....	104

2.6.5. Generation of artificially methylated and un-methylated genomic DNA.....	104
2.6.6. DNA gel electrophoresis.....	106
2.6.7. Purification of DNA samples and sequencing.....	106
2.7. Confirmation of the results of the dab oligonucleotide microarray using real-time PCR.....	106
2.7.1. Primer design, validation and product sequencing.....	107
2.7.2. Real-time PCR.....	109
2.8. Targeted quantification of 12 metabolites related to the one-carbon cycle via liquid chromatography (LC)-triple stage quadrupole (TSQ) tandem mass spectrometry.....	109
2.9. Statistical analysis.....	119
2.9.1. General statistical approaches used throughout this thesis.....	119
2.9.2. Statistical analysis of microarray data.....	119
2.9.3. Ingenuity Pathway Analysis of the microarray data.....	120
2.9.4. Principal component analysis and hierarchical clustering.....	120
Chapter 3: Comprehensive profiling of zebrafish hepatic proximal promoter CpG island methylation and its modification during chemical carcinogenesis.....	122
3.1. Introduction.....	123
3.2. Overview of the experimental approach (Methodological details are described in Chapter 2).....	126
3.3. Results.....	128
3.3.1. Global measurement of genomic DNA methylation.....	128
3.3.1.1. Establishing a method for measuring global genomic DNA methylation levels.....	128
3.3.1.2. Comparison of global genomic DNA methylation in fish and mammals using HPLC.....	130
3.3.2. DNA methylation at gene levels in zebrafish liver tumours.....	132
3.3.2.1. Unbiased enrichment of methylated DNA using methylated-DNA immunoprecipitation (MeDIP).....	132
3.3.2.2. Design of the zebrafish CGI tiling microarray and comprehensive mapping of adult zebrafish CGI methylation.....	136
3.3.2.3. DNA methylation analysis of zebrafish hepatocellular carcinoma and comparison to gene expression.....	143
3.3.2.4. Principal components analysis.....	150
3.3.3. Confirmation of the CGI tiling microarray data using bisulfite sequencing PCR.....	152
3.3.4. Gene expression and DNA methylation.....	155
3.4. Discussion.....	160
3.4.1. Overall DNA methylation levels in fish.....	160
3.4.2. DNA Methylation alterations at gene level in zebrafish HCC samples.....	161
3.4.3. Confirmation of the MeDIP-tiling microarray data using Bisulfite sequencing PCR.....	166
3.4.4. Comparison of gene expression and DNA methylation data.....	166
3.4.5. Comparison between DNA methylation profiles of HCC and healthy liver in zebrafish and in human.....	168

3.4.6. Conclusions.....	170
Chapter 4: DNA methylation in liver tumourigenesis in dab (<i>Limanda limanda</i>) from the environment.....	171
4.1. Introduction.....	172
4.2. Overview of the experimental approach (Methodological details are described in Chapter 2).....	176
4.3. Results.....	178
4.3.1. Global measurement of genomic DNA methylation.....	178
4.3.2. DNA methylation at gene levels in dab liver.....	180
4.3.2.1. Pilot study using MeDIP coupled with flounder cDNA microarray.....	180
4.3.2.1.1. Principal components analysis.....	181
4.3.2.2. Comprehensive DNA methylation profiling of dab HCA tumours.....	185
4.3.2.2.1. <i>De novo</i> high-throughput sequencing analysis of MeDIP DNA.....	185
4.3.2.2.2. Ingenuity Pathway Analysis.....	187
4.3.2.2.3. Confirmation of the MeDIP <i>de novo</i> high-throughput sequencing data using BSP and comparability of the data to additional individuals.....	191
4.3.2.2.4. Comparison of the BSP data.....	194
4.3.2.3. Transcription profiling of dab tumours.....	201
4.3.2.3.1. Design of 8x15k gene expression microarray and gene expression analysis of dab hepatocellular adenoma.....	201
4.3.2.3.2. Principal components analysis.....	202
4.3.2.3.3. Hierarchical clustering.....	202
4.3.2.3.4. Ingenuity Pathway Analysis.....	207
4.3.2.3.5. Confirmation of the gene expression data using RT-PCR.....	212
4.4. Discussion.....	214
4.4.1. Global DNA methylation.....	214
4.4.2. Pilot study indicated change in the methylation of DNA samples extracted from HCA samples compared to ST at gene level.....	215
4.4.3. Overview of the DNA methylation changes.....	216
4.4.4. DNA methylation and transcriptional changes associated with tumourigenesis in dab HCA.....	217
4.4.5. Methionine cycle and DNA methylation.....	218
4.4.6. A link between environmental contaminants, changes in DNA methylation and transcription, and dab liver tumours.....	221
4.4.7. Hypothesis.....	225
4.4.8. Conclusions.....	227
Chapter 5: Changes to the one-carbon cycle in dab liver tumours.....	228
5.1. Introduction.....	229
5.2. Overview of the experimental approach (Methodological details are described in Chapter 2).....	232
5.3. Results.....	234
5.3.1. Targeted quantification of 12 metabolites involved in the one-carbon cycle using LC-TSQ.....	234

5.3.2. Principal components analysis.....	240
5.3.3. Change in the expression and DNA methylation levels of the genes involved in the one-carbon cycle.....	242
5.4. Discussion.....	245
5.4.1. <i>S</i> -adenosylmethionine (SAM), <i>S</i> -adenosylhomocysteine (SAH) and DNA methyltransferases (DNMTs).....	245
5.4.2. The inhibitory role of homocysteine and its association with SAH.....	248
5.4.3. Alterations in the levels of primary nutrient methyl donors; choline, folate and methionine.....	249
5.4.3.1. Methionine and choline.....	250
5.4.3.2. Folate.....	251
5.4.4. Production of SAM and the competition between DNMTs, GAMT and PEMT for consumption of SAM.....	252
5.4.4. Conclusions.....	255
Chapter 6: General discussion and future work.....	257
6.1. General discussion.....	258
6.2. Main findings on global DNA methylation.....	259
6.3. Zebrafish and dab DNA methylation studies.....	260
6.4. Hypotheses for formation of tumours in dab liver.....	261
6.5. The importance of ST in this study and the impact of the epigenetic progenitor model of tumourigenesis.....	264
6.6. Gaps in knowledge and future work.....	266
6.7. Concluding remarks.....	269
Chapter 7: References.....	270
Chapter 8: Appendix.....	296
8.1. List of publications.....	297
8.2. List of abstracts for oral presentations.....	297
8.3. List of abstracts for poster presentations.....	298

“Additional files are saved on the enclosed CD”

List of Figures

Figure 1.1. The three key elements of epigenetic mechanisms (DNA methylation, post-transcriptional histone modifications and some forms of RNA such as small interfering RNA and microRNA) complement, stabilise and interact with each other to regulate transcription of genes.....	3
Figure 1.2. DNA methylation. Transfer of a methyl group from S-adenosylmethionine (SAM) via DNA methyltransferases to cytosine results in formation of 5-methylcytosine and S-adenosylhomocysteine (SAH).....	4
Figure 1.3. <i>De novo</i> and maintenance DNA methylation.....	5
Figure 1.4. The two DNA methylation reprogramming events during early embryogenesis and gonadal sex determination.....	7
Figure 1.5. MeCP2 structure and transcription suppression complex.....	16
Figure 1.6. Potential mechanisms of suppression of transcription via MBD containing MeCP2.....	17
Figure 1.7. Chromatin structures of active and inactive promoters.....	20
Figure 1.8. A schematic representation of the key DNA methylation and histone modifications that occur in the promoter region of a transcriptionally active gene.....	21
Figure 1.9. The link between DNA methylation, histone modifications and chromatin remodelling.....	22
Figure 1.10. The one-carbon cycle.....	26
Figure 1.11. Biological significance of DNA methylation in normal and tumour cells..	32
Figure 1.12. Demethylation/remethylation model for site specific hypomethylation in tumours.....	35
Figure 1.13. The viable yellow agouti (A^{vy}) mouse model.....	42
Figure 1.14. The interactions between environmental factors and the various “-omes”, including the epigenome.....	45
Figure 1.15. Multistage process of carcinogenesis.....	47
Figure 1.16. Epigenetic progenitor model of tumourigenesis.....	50
Figure 1.17. An image of the flatfish dab (<i>Limanda limanda</i>).....	56
Figure 2.1. Histopathology images of zebrafish HCC and healthy zebrafish liver.....	61
Figure 2.2. Confirmation of the type of dab liver lesion.....	67-68
Figure 2.3. Overview of the dab MeDIP-flounder cDNA microarray experiment.....	79
Figure 2.4. Scanned image of a flounder cDNA microarray from MeDIP-cDNA experiment.....	82
Figure 2.5. Agilent chamber.....	85
Figure 2.6. Zebrafish 4x44k tiling microarray.....	86
Figure 2.7. Overview of the MeDIP, HTS, assembly of contigs, design of the dab microarray and microarray experimental procedure.....	88

Figure 2.8. Equipment required for assessment of RNA integrity with Agilent 2100 Bioanalyzer.....	93
Figure 2.9. 2100 Bioanalyzer traces.....	94
Figure 2.10. Bisulfite sequencing PCR.....	97
Figure 2.11. Methprimer programme used for identification of CGIs and design of bisulfite sequencing PCR primers.....	99
Figure 2.12. Schematic representation of the multi-step gradient used for liquid chromatography.....	112
Figure 3.1. Flowchart of the experimental procedures used in Chapter 3.....	127
Figure 3.2. Comparison of HPLC, LC-MS/MS and ELISA methods for measurement of global levels of methylation in DNA.....	129
Figure 3.3.A. Determination of the retention times of five standard mononucleotides using HPLC. B. Spiked standard mixture with uridine monophosphate (UMP) for establishing the retention time of UMP for monitoring RNA contamination.....	131
Figure 3.4. Measurement of the global percentage of methylated cytosine in calf thymus, zebrafish, flounder and dab livers using HPLC.....	132
Figure 3.5. Direct bisulfite sequencing of <i>glutathione S-transferase P1</i> gene (<i>gstp1</i>)...	134
Figure 3.6. Direct bisulfite sequencing of <i>no tail</i> gene (<i>ntla</i>).....	135
Figure 3.7. Agarose gel electrophoresis image of methylated-DNA immunoprecipitated (MeDIP) bisulfite treated <i>ntla</i> and <i>gstp1</i> genes.....	136
Figure 3.8. Chromosomal mapping of the probes.....	138
Figure 3.9. Generation of artificially methylated genomic DNA.....	139
Figure 3.10. Chromosomal mapping of average methylation levels in tumour samples compared to control samples.....	140
Figure 3.11. Selected gene ontology (GO) terms significantly over-represented in the genes with DNA methylation levels 2-fold below the median level in zebrafish healthy liver i.e. genes with lower methylation levels (FDR< 5%).....	141
Figure 3.12. Gene ontology (GO) terms significantly over-represented in the genes with DNA methylation levels 2-fold above the median level in healthy zebrafish liver i.e. genes with higher methylation levels (FDR< 5%).....	142
Figure 3.13. Biological functions enriched among genes with hypermethylated levels in HCC samples compared to control samples.....	144
Figure 3.14. Biological functions enriched among genes with hypomethylated levels in HCC samples compared to control samples.....	145
Figure 3.15. Biological network of genes linked to the canonical pathway “molecular mechanisms of cancer” that were hypomethylated (>1.5-fold change, P<0.05) in zebrafish hepatocellular carcinoma compared to healthy liver.....	149
Figure 3.16. Principal component analysis (PCA) scores plot of DNA methylation data.....	151
Figure 3.17. Confirmation of the CGI tiling microarray using bisulfite sequencing PCR.....	153

Figure 3.18. Methylation levels in tumour and control for individual CpG sites for <i>coro2ba</i> , <i>igfbp1b</i> and <i>angptl3</i> genes.....	154
Figure 3.19. Ingenuity network predicted for genes that were both hypomethylated and had increased expression levels in zebrafish HCC (shaded red).....	159
Figure 4. 1. Flowchart of the experimental procedures used in Chapter 4.....	177
Figure 4.2. Measurement of global percentage of methylated cytosine.....	179
Figure 4.3. Principal component analysis (PCA) scores plot for all the genes that were differentially methylated in the HCA samples compared to ST in MeDIP-cDNA microarray data.....	184
Figure 4.4. Identification of differentially methylated contigs in dab HCA compared to surrounding tissue.....	187
Figure 4.5. Biological functions enriched among genes with altered methylation levels (hypo- and hyper-methylated) in HCA. IPA was used to group the genes with altered methylation levels based on biological functions (>1.5-fold change, FDR <5%).....	188
Figure 4.6. Direct bisulfite sequencing PCR data for A. <i>protocadherin 1 gamma 22</i> (<i>pcdhlg22</i> , -2.08), B. <i>microtubule-associated protein 1aa</i> (<i>map1aa</i> , -7.69), C. an unidentified sequence (+4.17), D. <i>tubulin tyrosine ligase-like member 7</i> (<i>ttl7</i> , +3.05) and E. <i>nidogen 1</i> (<i>nid1</i> , +2.01).....	192
Figure 4.7. BSP was used to compare the methylation levels of <i>protocadherin 1 gamma 22</i> , <i>microtubule-associated protein 1aa</i> , an unidentified sequence, <i>tubulin tyrosine ligase-like member 7</i> and <i>nidogen 1</i> in dab collected from five different sampling sites in the Irish Sea and the Bristol Channel.....	195
Figure 4.8. Principal component analysis (PCA) scores plot for all the genes that responded to the disease status of dab tissue samples. These genes were identified by the gene expression microarrays generated for HCA and ST.....	203
Figure 4.9. PCA scores plot for all the genes that responded to the disease status of dab tissue samples. These genes were identified by the gene expression microarrays generated for HCA and healthy tissue.....	204
Figure 4.10. PCA scores plot for all the genes that responded to the disease status of dab tissue samples. These genes were identified by the gene expression microarrays generated for HCA, ST and healthy tissue.....	205
Figure 4.11. Hierarchical clustering tree generated using Pearson correlation showing relationships between differentially transcribed genes and disease status (HCA and ST).....	206
Figure 4.12. Biological functions enriched among genes with altered transcription levels (over- and under-expressed) in HCA compared to ST.....	208
Figure 4.13. Biological functions enriched among genes with altered transcription levels (over- and under-expressed) in HCA compared to healthy dab liver.....	209
Figure 4.14. Biological network of genes linked to the canonical pathway “molecular mechanisms of cancer” with altered gene expression levels in dab HCA compared to ST (>1.5-fold change, FDR <5%).....	211
Figure 4.15. The methionine cycle. An illustration of the generation of the universal	220

methyl donor, SAM from methionine.....	
Figure 4.16. DNA methylation and transcription changes in genomic and non-genomic estrogenic pathway in HCA samples.....	224
Figure 4.17. A possible mechanism of tumourigenesis in dab.....	226
Figure 5.1. A schematic representation of an ion source coupled to triple quadrupole tandem mass spectrometer.....	231
Figure 5. 2. Flowchart of the experimental procedures used in Chapter 5.....	233
Figure 5.3. Example of a MS/MS spectrum collected from the specified 13 parent ions and product ions arising from the fragmentation of the parent ions.....	235
Figure 5.4. Quantification of the 10 metabolites of interest in hepatocellular adenoma (HCA), surrounding tissue (ST) and healthy dab liver samples.....	236 -239
Figure 5.5. Quantification of SAM/SAH ratio in HCA, ST and healthy dab liver samples.....	240
Figure 5.6. Principal component analysis (PCA) scores plot for the 10 metabolites that were investigated in the HCA, ST and healthy dab liver samples using LC-TSQ.....	241
Figure 5.7. Overview of the changes observed in dab HCA, ST and healthy dab liver samples as investigated in Chapter 5.....	254 -255
Figure 6.1. The complex network of epigenetic and non-epigenetic factors involved in the regulation of responses to environmental factors.....	268

List of Tables

Table 1.1. Studies investigating alterations in DNA methylation following treatment with environmentally relevant chemicals in aquatic species.....	11-12
Table 1.2. The methylation of the N-tails of histones is used as an epigenetic signature for predicting the transcription status of the genes.....	20
Table 1.3. Some of the identified hypermethylated genes in human cancers associated with key pathways in tumourigenesis.....	37
Table 2.1. Sample information. Cefas code, location of capture, sex and associated liver lesions of dab.....	63-65
Table 2.2. Bisulfite sequencing PCR primers used in zebrafish study (Chapter 3)...	100-102
Table 2.3. Bisulfite sequencing PCR primers used in dab study (Chapter 4).....	103
Table 2.4. Primer sequences for real-time PCR used for validation of the dab gene expression microarray results.....	108
Table 2.5. The multi-step gradient used for liquid chromatography.....	112
Table 2.6. The collision energy, S-lens value or stacked ring ion guide (SRIG) value (this value represents the radio frequency (RF) that efficiently captures and focuses the ions into a tight beam for transport of the ions from the source to the mass analyser), parent (precursor) masses, product masses, positive or negative ion mode and the chemical formula used for detection of the 13 compounds of interest are shown in this table.....	113-118
Table 3.1. Subsection of the genes with altered DNA methylation levels in HCC samples compared to control samples as measured by CGI-tiling microarray.....	146-148
Table 3.2. Genes with both significantly altered gene expression levels and DNA methylation levels (>1.5-fold change, $p < 0.05$).....	157-158
Table 4.1. Subsection of the identified clones with potentially altered methylation levels in dab HCA samples compared to ST determined by MeDIP-cDNA flounder microarray.....	182-183
Table 4.2. CpG island discovery.....	186
Table 4.3. Subsection of the genes with altered methylation associated with development of tumours (FDR <5%) in HCA compared to ST.....	189-190
Table 4.4. Subsection of the genes with altered gene expression in HCA compared to ST.....	210
Table 4.5. Comparison of fold change in gene expression for the selected genes by RT-PCR and microarray analysis.....	213
Table 5.1. Genes related to one-carbon pathway statistically significantly altered in expression between in the three categorise investigated (HCA, ST and H).....	244
Table 5.2. Genes with altered DNA methylation (> 1.5-fold enrichment) involved in the one-carbon cycle as quantified by <i>de novo</i> high throughput sequencing of 5-methylcytosine immunoprecipitated DNA.....	244

Abbreviations

5-AC	5-Azacytidine
5AdC	5-Aza-2' deoxycytidine
AFB1	Aflatoxin B1
AID	Activation-induced cytosine deaminase
AKT1	V-akt murine thymoma viral oncogene homolog 1
A(M/D/T)P	Adenosine (mono/di/tri)phosphate
ANGPTL3	Angiopoietin-like 3
APC	Adenomatous polyposis coli
API	Atmospheric pressure ionisation
As	Arsenic
ATF-2	Activating transcription factor 2
A ^{vy}	Viable yellow agouti allele
BAG5	BCL2-associated athanogene 5
Bcl-2	B-cell lymphoma 2
BCR-ABL fusion	Break cluster region-V-abl Abelson murine leukaemia viral oncogene homolog 1 fusion
BGI	Beijing Genomic Institute
B-H	Benjamini-Hochberg
BMHT	Betaine-homocysteine methyltransferase
BRCA1	Breast cancer type 1, early onset
BSP	Bisulfite sequencing PCR
BW	Body weight
CAR	Constitutive androstane receptor
CBS	Cystathionine β -synthase
CCND2	Cyclin D2
CCNDBP1	Cyclin D binding protein-1
Cd	Cadmium
CDH1	Cadherin 1
CDK8	Cell division protein kinase 8
CDKN2(A/B)	Cyclin-dependent kinase inhibitor 2(A/B)
(c)DNA	(complementary) Deoxyribonucleic acid
CDP-choline	Cytidine diphosphate-choline
Cefas	Centre for environment, fisheries and aquaculture science
Cfp1	CXXC finger protein 1
CG genes	Cancer-germline genes
CGIs	CpG islands
CGL	Cystathionine γ -lyase
CK	Choline kinase
CMS	Cytosine 5-methylenesulfonate
C(M/T)P	Cytidine (mono/tri)phosphate

CORO2BA	Coronin, actin binding protein 2BA
CPT	Cholinephosphotransferase
CSEMP	Clean seas environmental monitoring programme
CT	Threshold cycle
CTCF	CCCTC-binding factor
CTS	Cathepsin
CX43	Connexin-43
Cy(3/5)- dC(U)TP	Cyanine-(3/5) labelled deoxycytidine(uridine) triphosphate
CYP	Cytochrome P450
DAG	Diacylglycerol
d(A/C/G/T)MP	2'-deoxy(adenosine/cytidine/guanosine/thymidine) 5'-monophosphate
DAPK1	Death-associated protein kinase 1
DEFRA	Department for environment food and rural affairs
DES	Diethylstilbestrol
DMBA	7, 12-Dimethylbenz[α]anthracene
DMSO	Dimethyl sulfoxide
DNase 1	Deoxyribonuclease 1
DNMT	DNA (cytosine-5-)-methyltransferase
DNMT3L	DNA (cytosine-5-)-methyltransferase 3-like
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
DVL2	Dishevelled, dsh homolog 2
E	Embryonic day
E2	17 β -Estradiol
EAC	Ecotoxicological assessment criteria
EDCs	Endocrine disrupting chemicals
EDTA	Ethylenediaminetetraacetic acid
EE	17 α -Ethinylestradiol
E-ER	Estrogen-estrogen receptor
ELISA	Enzyme-linked immunosorbent assay
ENO2	Enolase 2
ER	Estrogen receptor
ERCC6	Excision repair cross-complementing rodent repair deficiency, complementation group 6
ERE	Estrogen response element
ERK	Extracellular-signal-regulated kinase
EST	Expressed sequence tag
ETS-TF	E-twenty six transcription factor
F _(0/1/2/3)	(parental/first/second/third) Filial generation
FCA	Foci of cellular alterations
FDR	False discovery rate

FE	Feature extraction
GAA	Guanidinoacetate
GAMT	Guanidinoacetate methyltransferase
GATA(4/5)	GATA binding protein (4/5)
GENIPOL	Genomic tools for biomonitoring of pollutant coastal impact
GNMT	Glycine <i>N</i> -methyltransferase
GO	Gene ontology
GSTP1	Glutathione <i>S</i> -transferase P1
HAT	Histone acetyltransferase
HBCD	Hexabromocyclododecane
HBV	Hepatitis B virus
HCA	Hepatocellular adenoma
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HDAC	Histone deacetylase
HIF-1	Hypoxia-inducible factor-1
H3(4)K9(27/36/20/4/79)me3	Trimethylation of lysine 9 (27/36/20/4/79) of N-terminal tail of H3 (4)
5-hmC	Hydroxymethylcytosine
HP1	Heterochromatin protein 1
hpf	Hour post-fertilisation
HPLC	High-performance liquid chromatography
HSP90B1	Heat shock protein 90- β member 1
HTS	High-throughput sequencing
HydVac	Hydropic vacuolation
IAP	Intracisternal A-particle
ICES	International council for the exploration of the seas
ICR	Imprinting control region
IGF2	Insulin growth factor 2
IGFBP	Insulin-like growth factor binding protein
IGF1R	Insulin growth factor-1 receptor
IPs	Inhibitory proteins
IPA	Ingenuity pathway analysis
Kb	Kilobase pair
Kdm2a	Lysine-specific demethylase 2A
LCMD	Laser capture microdissection
LOI	Loss of imprinting
Lsh	Lymphoid specific helicase
Lys/K	Lysine
MAGE-A1	Melanoma antigen A1
MAP1AA	Microtubule-associated protein 1AA
MAPK	Mitogen-activated protein kinase

MAT	Methionine adenosyltransferase
MBD	Methyl-CpG binding domain
5mC	5-Methylcytosine
5mdCMP	5-Methyl deoxycytidine 5'-monophosphate
ME	Molar equivalent
MeCP	Methyl-CpG binding protein
MeDIP	Methylated-DNA immunoprecipitation
5-methyl THF	5-Methyl tetrahydrofolate
MGMT	O-6-methylguanine-DNA methyltransferase
MLH1	MutL homolog 1
MMP14	Matrix metalloproteinase 14
MRM	Multiple reaction monitoring
MS	Methionine synthase
MSH(2/6)	MutS homolog (2/6)
MS/MS	Tandem mass spectrometry
MTHFR	Methylenetetrahydrofolate reductase
MYCB	Myelocytomatosis oncogene B
MZ	Monozygotic
M/Z	Mass to charge ratio
NID1	Nidogen 1
nr	Non-redundant
NTLA	No tail a
NuRD	Nucleosome remodelling and histone deacetylase
ONC	Oncogene
P53	Protein 53 or tumour protein 53
PAH	Polycyclic aromatic hydrocarbons
Pb	Lead
PB	Phenobarbital
PBDE	Polybrominated diphenyl ether
PC	Phosphatidylcholine
PC	Principal component
PCA	Principal components analysis
PCB	Polychlorinated biphenyl
PCDH1G22	Protocadherin 1 gamma 22
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
Pcytl	Phosphocholine cytidyltransferase 1
PE	Phosphatidylethanolamine
PEMT	Phosphatidylethanolamine <i>N</i> -methyltransferase
PPi	Pyrophosphate
ppm	Parts per million
q (1/2/3)	Quadrupole (1/2/3)

RAB2A	RAB2A, member of RAS oncogene family
RAP2C	RAP2C, member of RAS oncogene family
RB	Retinoblastoma
RF	Radio frequency
(c/m/r/t)RNA	(complementary/messenger/ribosomal/transfer) Ribonucleic acid
RNAi	Ribonucleic acid interference
RNAP	Ribonucleic acid polymerase
RPM	Revolutions per minute
RT-PCR	Real time-polymerase chain reaction
SAH	<i>S</i> -adenosylhomocysteine
SAHH	<i>S</i> -adenosylhomocysteine hydrolase
SAM	<i>S</i> -adenosylmethionine
SAM dependent	<i>S</i> -adenosylmethionine dependent methyltransferase
MT	
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
Set1(A/B)	Set domain containing 1(A/B)
SFRP2	Secreted frizzled-related protein 2
SNTB2	Syntrophin beta 2
SOAP	Short oligonucleotide analysis package
Sp1	Specificity protein 1
SRIG	Stacked ring ion guide
SSC	Sodium citrate sodium chloride
ST	Surrounding tissue
ST14	Suppression of tumourigenesis 14
STAT5	Signal transducer and activator of transcription 5
TBE buffer	Tris-borate-EDTA buffer
TBT	Tributyltin
TE buffer	Tris-EDTA buffer
TET1	Methylcytosine dioxygenase ten-eleven translocation 1
TF	Transcription factor
TFIIB	Transcription factor IIB
THF	Tetrahydrofolate
TIMP2	Tissue inhibitor of metalloproteinase 2
TPG	Tumour progenitor gene
TPT	Triphenyltin
TRD	Transcription repression domain
TRDMT1	tRNA aspartic acid methyltransferase 1
Tris	(Hydroxymethyl) methylamine
TSG	Tumour suppressor gene
TSQ	Triple stage quadrupole tandem mass spectrometry

TSS	Transcription start site
TTLL7	Tubulin tyrosine ligase-like member 7
TTP	Tristetraprolin
TTS	Transcription termination site
UGDH	UDP-glucose 6-dehydrogenase
U(M/D)P	Uridine 5'-(mono/di)phosphate
uPA	Urokinase-type plasminogen activator
VCZ	Vinclozolin
VGT	Vitellogenin
X _{a/i}	Active/inactive X chromosome
XIC	X inactivation centre
Xist	X (inactive)-specific gene

Chapter 1

General introduction

The information presented in this chapter contributed towards a review article.

Mirbahai et al., (2011) Epigenetic memory: Its significance in regulation of responses to toxicants in aquatic species and potential as a marker of life-time exposures, submitted to *Environmental Science & Technology*.

1.1. General introduction

The involvement of epigenetic mechanisms such as DNA methylation, although well established in human tumourigenesis, is relatively under-studied in tumours of aquatic species used in environmental monitoring or in the modeling of human disease.

Assessment of the disease status of the flatfish dab (*Limanda limanda*) is used in the UK Clean Seas Environmental Monitoring Programme (CSEMP), as an indicator of biological effects of contaminants in the marine environment. At some UK offshore sites the prevalence of liver tumours in dab exceeds 20% of sampled individuals (NMMP, second report, 2004; Stentiford et al., 2009; Small et al., 2010; Southam et al., 2008; Lyons et al., 2006); however, little research has been conducted to investigate the underlying mechanisms and, in particular, possible involvement of epigenetic factors. In addition to this, there is a great need for characterisation of DNA methylation changes in emerging human disease models, such as zebrafish (*Danio rerio*). Recent studies comparing hepatic gene expression in human and zebrafish tumours demonstrated conservation of gene expression profiles between these two species at different stages of tumour aggressiveness (Lam et al., 2006). However, the contribution of altered DNA methylation to such changes in zebrafish is unknown.

In this thesis, state of the art techniques were used to detect differences in the DNA methylation profiles of fish liver tumours compared to healthy fish livers. These approaches also helped to overcome the technical challenges encountered when working with unsequenced species. Combinations of ‘omic’ and gene/metabolite specific techniques at various levels of methylome, transcriptome and metabolome have been applied to determine the role of DNA methylation as an epigenetic factor in formation of liver tumours in dab and in chemically induced hepatocellular carcinoma in zebrafish. The ultimate aim of this project was to achieve a better understanding of the molecular mechanisms underlying environmental tumourigenesis.

1.2. Epigenetics

Epigenetics (above genetics, outside conventional genetics) is defined as meiotically and mitotically heritable changes in gene expression that are not due to alterations in primary DNA sequence (Egger et al., 2004; Jaenisch and Bird, 2003). Such modifications include

DNA methylation, post-transcriptional chemical modifications of the N-tail of histones, binding of non-histone chromatin proteins to DNA or histone modifications, small interfering RNA, nucleosome positioning and higher order organisation (Egger et al., 2004; Turner, 2007; Chan et al., 2008; Cui et al., 2010; Probst et al., 2009). Moreover, these modifications are not isolated events and are closely inter-linked by influencing chromatin structure at various levels and by further interactions with the genome (Figure 1.1) (Probst et al., 2009).

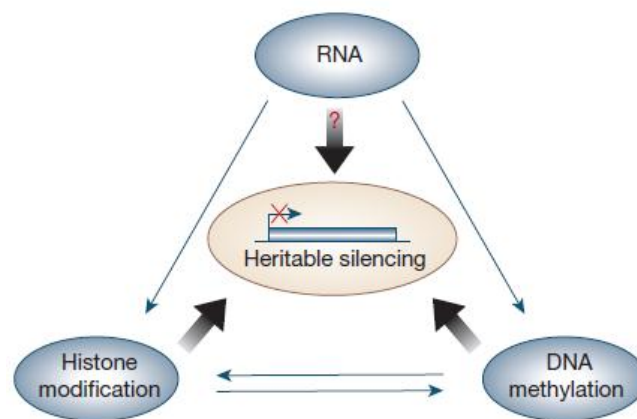


Figure 1.1. The three key elements of epigenetic mechanisms (DNA methylation, post-transcriptional histone modifications and some forms of RNA such as small interfering RNA and microRNA) complement, stabilise and interact with each other to regulate transcription of genes. Disruption of any of the elements of this interacting system may lead to inappropriate regulation of gene expression and development of a range of disorders (Egger et al., 2004).

Normal cells display finely tuned epigenomic equilibrium (Esteller and Herman, 2002). Disruption of this equilibrium leads to inappropriate regulation of transcription, hence giving rise to diseases and disruption of differentiation and developmental processes (Egger et al., 2004; Probst et al., 2009). Methylation of DNA at CpG dinucleotides is the most studied epigenetic modification (Esteller and Herman, 2002) and is the principal focus of this thesis, especially in regards to tumourigenesis in fish.

1.3. DNA methylation

DNA methylation is the most studied epigenetic mechanism associated with regulation of gene expression (Esteller, 2007; Pogribny, 2010). To methylate DNA, a methyl group is transferred from the universal methyl donor, *S*-adenosylmethionine (SAM), to the 5th carbon position of a cytosine base. This reaction is catalysed by a group of enzymes known as DNA methyltransferases (DNMTs) (Figure 1.2). In eukaryotes this modification is predominantly observed in cytosine bases located at CpG dinucleotide sequences (Pogribny, 2010; Gronbaek et al., 2007). In mammals CpG dinucleotides are not randomly distributed throughout the genome. In fact, higher frequency of CpG sites are found within the regions referred to as CpG islands (CGI). These sites have distinctive characteristics, including higher frequencies of CpG dinucleotides relative to the bulk genome, G+C percentages of greater than 50% and average lengths greater than 200bp (Gardiner- Garden and Frommer, 1987). Furthermore, CpG dinucleotide sites located in CGIs are mostly un-methylated and are detected within the promoter regions and first exons of 40_50% of mammalian genes. The remaining CpG sites located at the non-CGIs are mostly methylated in mammals. This means that approximately 50_70% of all CpG dinucleotides in the mammalian genome are methylated (Karp, 2000; Jones and Takai, 2001; Taylor et al., 2007; Suzuki and Bird, 2008; Fang et al., 2008; Illingworth and Bird, 2009; Pogribny, 2010).

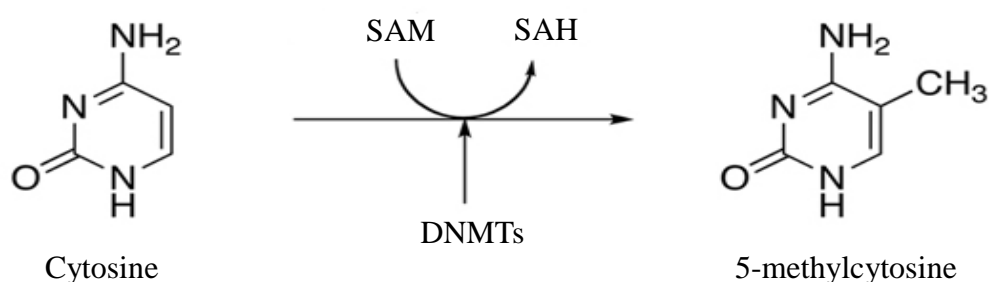


Figure 1.2. DNA methylation. Transfer of a methyl group from *S*-adenosylmethionine (SAM) via DNA methyltransferases to cytosine results in formation of 5-methylcytosine and *S*-adenosylhomocysteine (SAH).

1.3.1. DNA methyltransferases; *de novo* and maintenance DNA methylation

In mammals, four different DNA methyltransferase (DNMT) members have been identified corresponding to two distinct families. The DNMT3 family establishes the DNA methylation pattern in previously un-methylated DNA of stem cells and tumour cells. This is known as *de novo* methylation. DNMT1 is responsible for maintenance of the established DNA methylation pattern during cell replication in differentiated cells (Figure 1.3). DNMT1, cloned by Bestor et al., (1988), utilises the hemi-methylated DNA strand as a template for reproducing the exact methylation pattern in the newly synthesised DNA strand, thereby ensuring conservation and inheritance of the DNA methylation pattern in the daughter cells (Momparker and Bovenzi, 2000; Gronbaek et al., 2007; Lopez et al., 2009; Cheng and Blumenthal, 2008). The highest concentrations of DNMT1 are reported during S phase. During this stage, DNMT1 interacts with proliferating cell nuclear antigen (PCNA) and therefore, its localisation at the replication fork throughout the replication of the DNA is ensured (Jones and Liang, 2009).

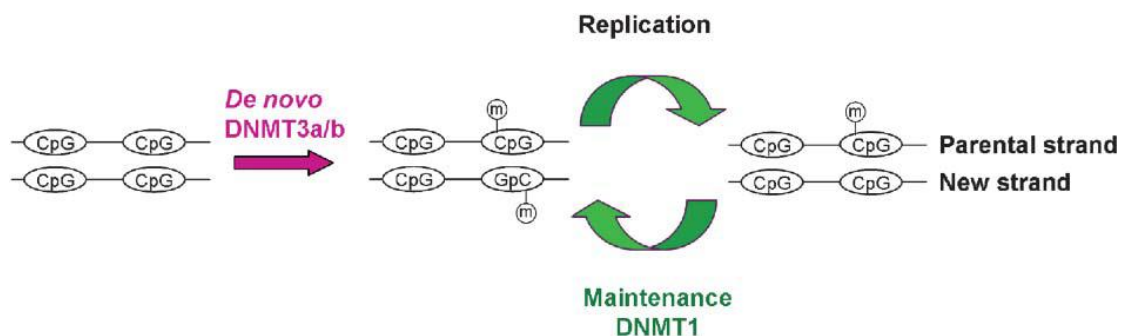


Figure 1.3. *De novo* and maintenance DNA methylation. Methylation of the un-methylated DNA in embryonic stem cells and cancer cells, known as *de novo* methylation, is mainly conducted by DNMT3a and 3b. Following each round of replication the maintenance methyltransferase, DNMT1, copies the DNA methylation pattern from the parental strand onto the newly synthesised strand. M: 5-methylcytosine. Adapted from Gronbaek et al., (2007).

Detection of cycles of DNA methylation and de-methylation with corresponding cycles of transcription in human cells provided the evidence that DNA methylation may be more plastic than previously believed (Kangaspeska et al., 2008). Indeed, at early stages of embryogenesis

and gametogenesis the DNA methylation pattern is largely erased, reaching its lowest level around implantation time of embryogenesis. Although the precise mechanisms of DNA de-methylation are not known, it has been suggested that DNA can either passively or actively undergo de-methylation (Herceg and Vaissiere, 2011). During passive de-methylation, 5-methylcytosine is removed in a replication-dependent manner by inhibition of DNMTs while active de-methylation depends upon enzymatic removal of 5-methylcytosine. During active de-methylation 5-methylcytosine is converted to thymidine. This can occur both spontaneously or it can be catalysed by activation-induced cytosine deaminase (AID) protein. The generated mismatched T:G can be recognised by the repair machinery and converted back to cytosine (Herceg and Vaissiere, 2011). Recently, it has been discovered that methylcytosine dioxygenase ten-eleven translocation 1 (TET1) is involved in regulation of DNA methylation. TET1 protein catalyses the conversion of 5-methylcytosine to 5-hydroxymethyl cytosine (5-hmC). 5-hmC sites are poorly recognised by DNMTs. This results in a passive replication-dependent loss of DNA methylation. Alternatively, 5-hmC can be recognised by the repair machinery and converted to cytosine (Tahiliani et al., 2009; Williams et al., 2011; Herceg and Vaissiere, 2011).

Shortly after, the erased DNA methylation pattern at implantation stage is re-established via *de novo* methyltransferases (Figure 1.4). *De novo* methylation is mediated by the two active members of DNMT3 family namely DNMT3a and DNMT3b, cloned by Okano et al., (1998), and one regulatory factor, DNMT3-like protein (DNMT3L). The latter is only identified in germ cells and lacks the conserved catalytic domain present in C-terminal of DNMT3a and DNMT3b, hence lacks any methyltransferase activity. Nevertheless, it has been shown that DNMT3L has a role as a *de novo* regulatory factor and DNMT3 stabiliser during establishment of DNA methylation patterns of sub-set of genes in both male and female germ cells (Cheng and Blumenthal, 2008; Goll and Bestor, 2005; Laird and Jaenisch, 1994). Although all members of DNMT3 family are mainly associated with *de novo* methylation, they have the same affinity towards both hemi- and un-methylated DNA. Therefore, in addition to their role in *de novo* methylation, they also assist in maintenance of the DNA methylation pattern (Gronbaek et al., 2007).

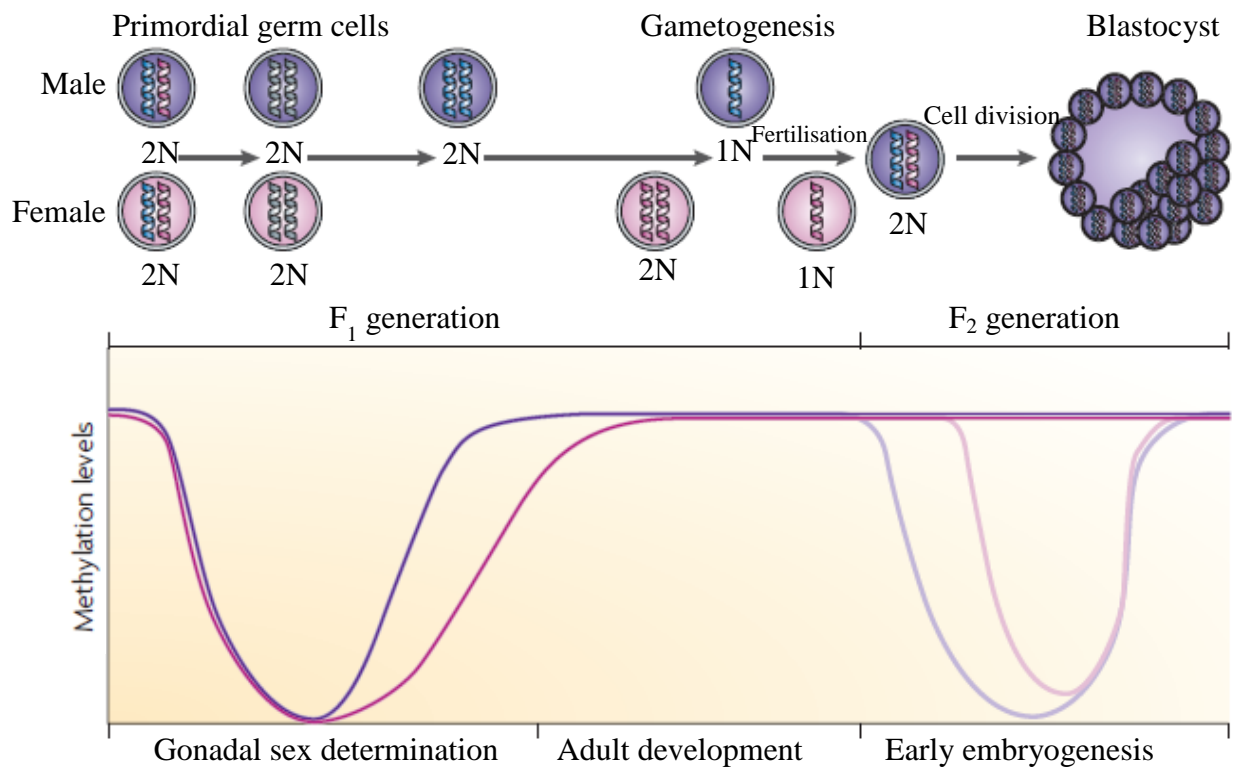


Figure 1.4. The two DNA methylation reprogramming events during early embryogenesis and gonadal sex determination. During gonadal sex determination, DNA methylation patterns of primordial germ cells are erased, then DNA methylation patterns are re-established in the developing male (purple line) and female (pink line) germ lines based on paternal and maternal-origin of genes (i.e. imprinting), respectively (imprinting is discussed in section 1.3.4.1). The re-methylation in female germ cells is established at a later stage than in male germ cells. Following fertilisation, the paternal genome (lighter purple line) and maternal genome (lighter pink line) undergo a second wave of genome-wide de-methylation, reaching lowest levels of methylation during implantation. During this round of de-methylation, parenteral imprinting markers are protected against de-methylation. Methylation levels are re-established in blastocysts (Jirtle and Skinner, 2007).

However, it is not clear how certain CGI sites are protected against methylation while other sites are methylated during *de novo* methylation. Two possible models have been proposed:

1. **Non-selective model:** In this model it is proposed that DNA is methylated at all CpG dinucleotides that are not protected against DNA methylation. CpG islands contain a binding site for transcription factor specificity protein 1 (Sp1). Removal of the Sp1 binding site results in methylation of CGIs. Therefore, it is possible that Sp1 binding to CGIs provides protection against DNA methylation (Jones and Takai, 2001).

2. Targeted model: This model is based upon the realisation of the complexity of the DNA methylation procedure. For DNA methylation to occur interaction of several factors, such as chromatin remodelling complexes, histone modifying enzymes and DNA methyltransferases are required. Thus absence of any of these components from a CpG site prevents the methylation procedure (Jones and Takai, 2001). For example, the absence of chromatin remodelling proteins at active regions of DNA prevents structural alterations of chromatin which otherwise would have allowed DNMTs to access and methylate DNA (chromatin remodelling is discussed in section 1.3.2) (Doerfler, 1983). Lsh (lymphoid specific helicase) is a member of the SNF2/helicase family with chromatin-remodelling properties. Interaction between this protein and DNMT3a, DNMT3b, heterochromatin protein 1 (HP1) and other unknown proteins leads to formation of a remodelling complex. This complex initiates *de novo* DNA methylation and formation of heterochromatin at these regions (Yan et al., 2003; Zhu et al., 2006). Disruption of Lsh function during embryogenesis has been linked to hypomethylation, alteration in histone H3 methylation and increase in histone H3 and H4 acetylation. Lsh only contributes to regulation of *de novo* DNA methylation and it appears not to be involved in DNMT1 activity during maintenance methylation (Zhu et al., 2006).

In addition, a third family of methyltransferases, DNMT2, with high sequence and structural homology to the previously described DNMTs has been identified. Due to conservation of the sequence of the catalytic motif in DNMT2, this protein was originally classified as DNA methyltransferase. However, rather than participating in epigenetic regulation of transcription by promoting hypermethylation of DNA sequences, DNMT2 mediates the methylation of the cytosine-38 residue in the anticodon loop of aspartic acid transfer RNA (tRNA). Hence to reflect its biological function the name of this protein was changed to tRNA aspartic acid methyltransferase 1 (TRDMT1). Although the molecular mechanisms and the biological significance of methylation of tRNA by DNMT2 are not entirely established, several studies in zebrafish (*Danio rerio*), *Drosophila* and mouse (*Mus musculus*) embryonic fibroblast cells have predicted a critical role for DNMT2 throughout the lifespan of the tested species. Indeed, morpholino knockdown of this gene in zebrafish results in lethal defects (Cheng and Blumenthal, 2008; Goll and Bestor, 2005; Bourchis et al., 2001; Schaefer et al., 2010). In addition, recently Schaefer et al., (2010) have identified two additional substrates of DNMT2: tRNA^{gly-GCC} and tRNA^{val-AAC}. They are both methylated at C38 residue located in the anticodon loop. DNMT2-mediated methylation protects the anticodon regions of tRNAs

against endonucleolytic cleavage induced under thermal and oxidative stress conditions. This protective role of methylation is mediated by methylation-induced structural alterations of the tRNA-substrates and by restricting access of endonuclease to the anti-codon regions.

DNA methylation, although not as well-studied as in humans, occurs in other non-mammalian species as demonstrated in zebrafish. Both DNA methyltransferase genes and DNA methylation reprogramming that occur during embryogenesis in humans have been detected in adult and zebrafish embryos, respectively (Mhanni and McGowan, 2004; MacKay et al., 2007; Martin et al., 1999). In total eight different DNA methyltransferases have been identified in zebrafish (*dnmt1-dnmt8*). Similar to human *DNMT3*, two splice variants for *dnmt5* and *6* have been cloned in zebrafish. However, only the biological functions of *dnmt1* and *dnmt7* are known. Dnmt1 protein is found abundantly from the time of fertilisation until 6 hours post-fertilisation (hpf). As transcription in the embryo starts from 3 hpf (1000 cell), this suggests that the origin of the Dnmt1 is maternal. The embryonic expression of *dnmt1* gene increases in the brain and neural tube at 16 hpf, and reaches its highest levels at 24 hpf in the eyes and central nervous system. It is thought that this particular enzyme may be responsible for maintenance of DNA methylation. In contrast, Dnmt7 activity has been associated with *de novo* methylation of specific genes, such as the *no tail (ntl)* gene. As Dnmt7 is detected as early as the 4-cell stage, this suggests an initial maternal origin for Dnmt7 (Shimoda, 2007).

Initially genome-wide embryonic DNA methylation reprogramming was un-detected during embryogenesis in zebrafish and thereby considered to be absent in zebrafish (Macleod et al., 1999). This raised the possibility of overestimating the evolutionary role of DNA methylation in differentiation and reprogramming. However recent studies revealed that the zebrafish, similar to humans, undergoes extensive DNA hypomethylation during early stages of development, reaching its lowest levels at cleavage stage (1.5 hpf), early blastula stage (2 hpf) and mid-blastula (3 hpf) stage with DNA re-methylation during late blastula stage (4 hpf). As this process occurs at very early stages of development of the embryo and over a short time-period (4 h), it had previously gone undetected (Mhanni and McGowan, 2004; Mackay et al., 2007; Martin et al., 1999). Therefore DNA methylation is not restricted to mammals. Indeed although limited, several studies have been published on the effect of common environmental contaminants on DNA methylation levels of aquatic species, either at gene level or at global

level. As environmental contaminants are the key determinants of the health of aquatic species, these studies and their main outcomes are summarised in Table 1.1.

In addition to zebrafish, DNA methylation has been detected throughout the evolutionary tree in arthropods, non-arthropod invertebrates and vertebrates (Bird, 1980). Furthermore, DNA methyltransferases are extremely well conserved between different species; thus, implying a significant role for DNA methylation throughout evolution (Lee et al., 2010).

However, it is important to mention that DNA methylation patterns and levels vary throughout the animal kingdom. For instance, DNA methyltransferases have not been detected in the nematode worm *Caenorhabditis elegans*. As a result its genome lacks any detectable methylated cytosines. The differences in the DNA methylation patterns within the animal kingdom suggest the potential for different functions for DNA methylation systems in some animals (Bird, 2002). Therefore, the role of DNA methylation cannot be generalised for all species.

Test species	Tested compound	Targeted organ	Investigated epigenetic marker	Main finding	Ref
Adult zebrafish (<i>Danio rerio</i>)	EE: 100 ng/L Exposure period: 14 days	Liver and brain	5' flanking region of <i>vtgI</i>	Hypomethylation of the <i>vtgI</i> gene in the liver corresponds with its higher expression following treatment with EE. Transcriptionally suppressed <i>vtgI</i> in the brain was methylated and its methylation and transcription did not respond to EE treatment.	Stromqvist et al., 2010
Three-spined stickleback (<i>Gasterosteus aculeatus</i>)	HBCD: 30 and 300 ng/L for 30 days, E2: 100 ng/L, 5AdC: 0.5 µg/g body weight. E2 and 5AdC exposures were for 22-23 days	HBCD: female liver, E2: Female and male liver and gonads, 5AdC: female and male gonads and liver	Global DNA methylation levels	Significant hypermethylation of genomic DNA in the male gonads following treatment with E2 and 5AdC with trends of hypermethylation in females. Significant global DNA hypomethylation of the female liver after treatment with 5AdC with trends of hypomethylation in males. Trends of hypomethylation in the female liver following exposure with E2 with no change detected in males. No change was detected in female livers following treatment with HBCD.	Aniagu et al., 2008
Goldfish (<i>Carassius auratus</i>)	Cu (100 µg/L), Zn (50 µg/L), Pb (20 µg/L), Cd (10 µg/L) and a mixture of all four metals. Exposure period: 48 h	Liver	Global DNA methylation levels	Significant DNA hypermethylation was detected in the liver. This effect was more pronounced for Cd and Pb compared to Cu and Zn. A concentration-dependent increase in overall DNA methylation levels were detected for the mixture.	Zhou et al., 2001
Japanese medaka (<i>Oryzias latipes</i>)	EE: 500 ng/L for 14 days	Liver, brain and gonads	Estrogen receptor gene and aromatase gene	Tissue and gender differences in methylation levels before and after treatment with EE. However, the CpG sites investigated were limited and may explain the lack of correlation between gene expression and DNA methylation	Contractor et al., 2004

				changes observed following treatment with EE.	
Bluegill sunfish (<i>Lepomis macrochirus</i>)	benzo[α]pyrene: 1 µg/L for 40 days	Liver	Global DNA methylation levels	Trends of global hypomethylation from day 2 of the exposure.	Shugart, 1990
Male false kelpfish (<i>Sebastiscus marmoratus</i>)	TBT: 1,10 and 100 ng/L, TPT: 1, 10 and 100 ng/L, mixture of both: 0.5 ng/L+0.5 ng/L and 5 ng/L + 5 ng/L, all exposures were conducted for 48 days	Liver	Global DNA methylation levels	Detection of significant dose-dependent DNA hypomethylation after exposure to chemicals.	Wang et al., 2009
Water flea (<i>Daphnia magna</i>)	Cd (180 µg/L) for 10 days and 4 days in the following generation	Whole organism	Two un-known methylated fragments	No change in methylation of the two fragments after exposure to Cd.	Vandeghechuchte et al., 2009a
Water flea (<i>Daphnia magna</i>)	Zn (388 µg/L), 21 day exposures	Whole organism	Global DNA methylation levels	Detection of significant DNA hypomethylation in the first generation (F ₁) after the treated group (F ₀). This change did not proceed to the next generation (F ₂)	Vandeghechuchte et al., 2009b; Vandeghechuchte et al., 2010a
Water flea (<i>Daphnia magna</i>)	5AC: 2.9 mg/L (F ₀), 2.3 mg/L (F ₁), Genistein: 4.7 mg/L (F ₀ –F ₂), vinclozolin: 0.54 mg/l (F ₀), 0.45 mg/l (F ₁), 0.18 mg/l (F ₂). 15–17 days post exposure	Whole organism	Global DNA methylation levels	Vinclozolin and 5-AC induced significant DNA hypomethylation in F ₀ groups. The effects of vinclozolin were not transferred to non-exposed F ₁ and F ₂ offspring. However, DNA hypomethylation induced by 5-AC in F ₀ group was transferred to non-exposed F ₁ and F ₂ offspring.	Vandeghechuchte et al., 2010b

Table 1.1. Studies investigating alterations in DNA methylation following treatment with environmentally relevant chemicals in aquatic species. Abbreviations; EE: 17α-ethinylestradiol, HBCD: hexabromocyclododecane, E2: 17β-estradiol, 5AdC: 5-aza-2' deoxycytidine, *vtg*: vitellogenin I, 5-AC: 5-azacytidine, TBT: tributyltin, TPT: triphenyltin, F_(0,1,2): (parental, first, second) filial generation.

1.3.2. DNA methylation and regulation of gene expression: Crosstalk between DNA methylation and histone modifications

DNA methylation alone, as an epigenetic mechanism, is not sufficient for regulation of transcription. Undeniably, this procedure is a complex series of events involving a range of epigenetic and non-epigenetic factors and chemical modifications of DNA and histones. These changes will lead to chromatin remodelling and subsequently alter DNA-protein interactions (Doerfler, 1983; Gronbaek et al., 2007; Li, 2002). The crosstalk between DNA methylation, histone modifications and chromatin remodelling proteins is required for reinforcing the outcome. However, the precise mechanisms, the order of events and the factors linking these epigenetic mechanisms are not entirely known (Butler et al., 2008).

1.3.2.1. DNA methylation as a gene silencing mechanism

Transcriptional silencing of genes via DNA methylation is divided into two broad categories: 1. DNA methylation and enhancer blocking activity of insulators. 2. Methyl-CpG binding proteins (MeCPs).

1. DNA methylation and enhancer blocking activity of insulator-binding proteins: Insulators are DNA sequence elements that have a significant role in regulation of transcription through several mechanisms, such as inhibition of access of enhancers to promoter regions of genes (Burgess-Beusse et al., 2002). A number of genes are regulated by the same enhancer elements located at a distance from the transcription sites of these genes. Inhibition of inappropriate activation of genes under the control of the same enhancer is achieved by blocking the access of the enhancer to the promoter region with insulator-binding proteins. Insulator-binding proteins recognise and bind to insulator DNA sequences and regulate transcription of the genes by positioning themselves between enhancer and the transcription start site of the genes (Bell et al., 1999; Hark et al., 2000; Jaenisch and Bird, 2003). Methylation of insulator DNA sequences inhibits binding of the insulator-binding proteins to these regions, thus allowing access of the enhancer to the promoter region and expression of the immediate gene. For example, the imprinted insulin growth factor 2 (*IGF2*) gene with an inactive maternal allele and active paternal allele is regulated by the insulator binding protein, CCCTC-binding factor (CTCF). CTCF recognises an insulator sequence element in the imprinting control region (ICR) prior to *IGF2* gene. The ICR region is un-methylated and

methylated in maternal and paternal alleles, respectively. This results in binding of CTCF to its recognition site, located between the enhancer and promoter region of the un-methylated maternal allele. Therefore, transcription is prevented by blocking access of the enhancer. In contrast, as the ICR region is methylated in the paternally inherited allele, binding of CTCF and blocking of the enhancer is prevented, thereby resulting in expression of paternal allele of *IGF2* gene (Hark et al., 2000).

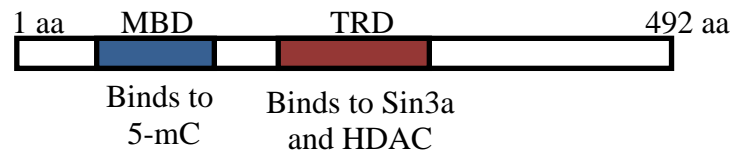
2. The methyl-CpG binding proteins (MeCPs). Methylation of the cytosine ring modifies the shape and the structure of the major groove, and this directly inhibits the binding of transcription factors (TFs) by preventing access to their binding sites (Counts and Goodman, 1995; Shi et al., 2003; Gronbaek et al., 2007; Momparler and Bovenzi, 2000; Li, 2002). However, the predominant mechanism of suppression of transcription via DNA methylation relies on selective recognition of methylated CpG dinucleotides by specific proteins that contain a methyl-CpG binding domain (MBD). These proteins include, MeCP2 and MBD1_4 (Li, 2002; Wade, 2001). Although the precise mechanisms and components of the complexes formed requires further investigation (Li, 2002), it is thought that MBD-containing proteins selectively recognise and bind to methylated cytosine regions, thereby promoting formation of a complex comprised of histone deacetylase (HDAC), histone methyltransferases and adenosine triphosphate (ATP)-dependent chromatin remodelling proteins. Removal of the acetyl groups from the N-tail of histones and remodelling of the chromatin results in the wrapping of the DNA around the core histones and formation of condensed transcriptionally inactive heterochromatin regions (Li, 2002; Wade, 2001). MeCP2 is the first identified and the best characterised member of the MBD family. It contains two functional domains, MBD and transcription repression domain (TRD), which promote formation of a suppressive complex (Figure 1.5.B). The MBD mediates binding of MeCP2 to methylated cytosine bases of the DNA while TRD provides a binding site for the co-repressor proteins known as Sin3a and HDACs (Figure 1.5.A) (Wade, 2001; Li, 2002; Hendrich et al., 2001). Figure 1.6, adapted from Wade (2001), demonstrates the potential mechanisms of transcription repression mediated by MBD containing MeCP2 protein. Although this model (indicated as model 1 in Figure 1.6) is certainly the predominant model of DNA methylation mediated transcription suppression, there are alternative HDAC-independent MBD mediated suppressive models that are not as well characterised, as shown in Figure 1.6. For example MBD can prevent binding of TF by directly occupying their binding sites (model 2), promoting remodelling of

chromatin architecture (model 3) or inhibiting transcription by binding to general transcription factors that are part of the RNA polymerase II pre-initiation complex, such as transcription factor IIB (TFIIB) (model 4). It is thought that the chosen mechanism is influenced by the type of cell, chromatin architecture and DNA sequence (Wade, 2001).

The function of the MBD1-3 transcription repression proteins is suppressed in the presence of HDACs inhibitors. This indicates the presence of a transcription repression domain (TRD) with HDAC-dependent activity in the structure of these proteins (Wade, 2001). Although the mammalian MBD3 protein contains a methyl-CpG binding domain, it lacks the capability of recognising methylated DNA. Therefore, the transcription repression ability of MBD3 is brought about by formation of a nucleosome remodelling and histone deacetylase (NuRD)-complex. This complex is comprised of histone binding proteins, HDAC and mi-2. The latter is an ATP-dependent chromatin remodelling protein which promotes transient changes in the architecture of chromatin. Furthermore, HDACs induce DNA condensation and suppress transcription through hypoacetylation of DNA (Hendrich et al., 2001; Li, 2002; Wade, 2001).

In comparison to MeCP2 and MBD3, the biochemistries of the MBD1 and MBD2 are not well characterised. However, as both contain TRD and MBD domains, it is thought that their transcription repression activities at methylated regions of DNA are due to recruitment of HDACs (Wade, 2001; Li, 2002). MBD2, similar to MBD3, interacts with the NuRD complex. The MBD2-NuRD complex, which was previously referred to as MeCP1 complex, also demonstrates chromatin remodelling properties. In contrast to the transcription suppression activity of MBD1-3, MBD4 is a DNA mismatch repair protein with DNA glycosylase activity (Li, 2002).

A.



B.

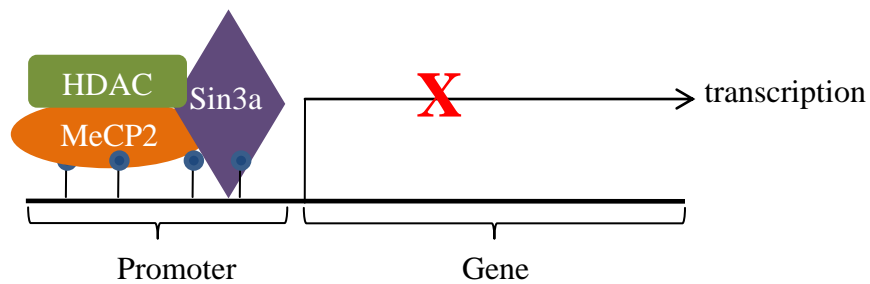
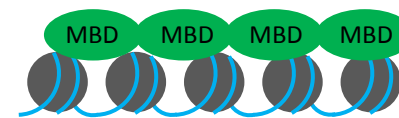


Figure 1.5. MeCP2 structure and transcription suppression complex. A. MeCP2 contains two functional domains: methyl-CpG binding domain (MBD), which recognises and binds to methylated cytosine and transcription repression domain (TRD) which can bind to co-repressor Sin3a and recruit histone deacetylases. B. MeCP2 selectively recognises methylated cytosines and promotes the formation of a complex, containing Sin3a co-repressor and HDAC. This will result in histone hypoacetylation and chromatin condensation.

Model 1. MBD proteins interact with HDAC to generate hypoacetylated, condensed chromatin



Model 2. MBD proteins coat methylated loci blocks regulatory regions of DNA



Model 3. MBD proteins alter local DNA and or chromatin architecture



Model 4. An MBD protein binds an essential transcription factor, preventing its function

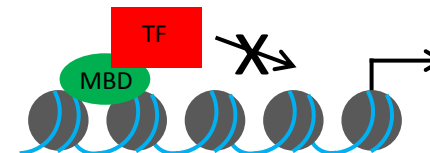


Figure 1.6. Potential mechanisms of transcription suppression via MBD containing MeCP2. Model 1. Binding of the MeCP2 to the methylated regions of DNA via its MBD binding domain results in recruitment of histone deacetylases (HDAC). HDACs cause the hypoacetylation of the chromatin N-tail and chromatin condensation. Model 2. Physical occupation of the transcription factor binding sites by MeCP2 inhibits activation of transcription. Model 3. MeCP2 can promote remodelling of chromatin structure and condensation. Model 4. Some transcription factors can bind to MeCP2 proteins thereby preventing their interaction with their regulatory regions. DNA: blue line, core histones: gray circles, MeCP2: is shown by its MBD domain presented by a green circle, HDAC: red circle, TF: red square. This figure was modified from Wade, 2001.

1.3.2.2. DNA methylation in relation to active transcription

Transcribed regions are associated with markers commonly indicative of active transcription status, such as, un-methylated promoter regions and certain post-transcriptional chemical modifications of the histones including acetylation. Transfer of the negatively charged acetyl group to the N-tails of the histones via histone acetyltransferases (HATs) results in neutralisation of the positive charge of the histones. This reduces the interaction between the negatively charged phosphate groups of DNA and positively charged N-tails of the histones, resulting in formation of euchromatin regions with relaxed DNA-protein conformations. These alterations allow access of transcription factors to the regulatory regions and initiate transcription (Figure 1.7) (Jones and Takai, 2001; Gronbeak et al., 2007). In tumourigenesis, alterations in methylation of the lysine (mono, di, tri) residues situated at the N-tails of the histones are as significant as changes in acetylation of these sites (Tate et al., 2010; Gronbaek et al., 2007). However, in contrast to acetylation, which is independent of the location of the affected amino acid residue, the influence of the methylation upon gene expression is highly dependent on the location of the lysine in the histone tail (Table 1.2) (Gronbaek et al., 2007). The crosstalk between histone modification and DNA methylation at the transcribed non-methylated CGI containing promoter regions is accomplished by a epigenetic marker known as, CXXC finger protein 1 (cfp1) (Thomson et al., 2010). This protein with its Cys-rich CXXC DNA binding domain selectively binds to methylation-free CpG dinucleotides (Thomson et al., 2010; Tate et al., 2010; Butler et al., 2008). This binding will further result in recruitment of the Set1A (Set domain containing 1A) and Set1B (Set domain containing 1B) histone H3K4 methyltransferase complex. Set1A and Set1B are both mammalian histone methyltransferases with non-overlapping, sub-nuclear localisation patterns. They exclusively bind to non-methylated, cfp1-attached, and transcriptionally active euchromatin regions and catalyse the methylation of H3K4, an epigenetic marker associated with transcriptionally active genes. H3K4 methylation results in subsequent interaction of the Set1A or Set1B with the RNA polymerase (RNAP) II C-terminal domain and initiation of transcription (Tate et al., 2010; Butler et al., 2008). In addition to the changes discussed, lysine-specific demethyltransferase 2A (KDM2A) catalyses the demethylation of H3K36 at these regions. Methylated H3K36 is a chemically modified histone, commonly associated with suppressed transcription (Figure 1.8) (Thomson et al., 2010).

Based on the mechanisms of suppression and activation of transcription described above, and as shown in Figure 1.7, it is clear that histone modifications, DNA methylation and ATP-dependent chromatin remodelling are interdependent processes. For example, in *Arabidopsis*, heterochromatin protein 1 (HP1) binds to methylated H3-K9 which is an epigenetic marker of inactive genes. This results in recruitment of *Arabidopsis* methyltransferase and initiation of methylation of CpNpG sequences (Li, 2002). However, establishing the order of events i.e. DNA methylation, promotion of histone modifications and chromatin remodelling and investigating the role of DNA methylation in regulating transcription, requires further investigation, especially in complex species such as mammals and plants. However, three possible scenarios for the order of interactions between different epigenetic factors have been proposed and are shown in Figure 1.9.

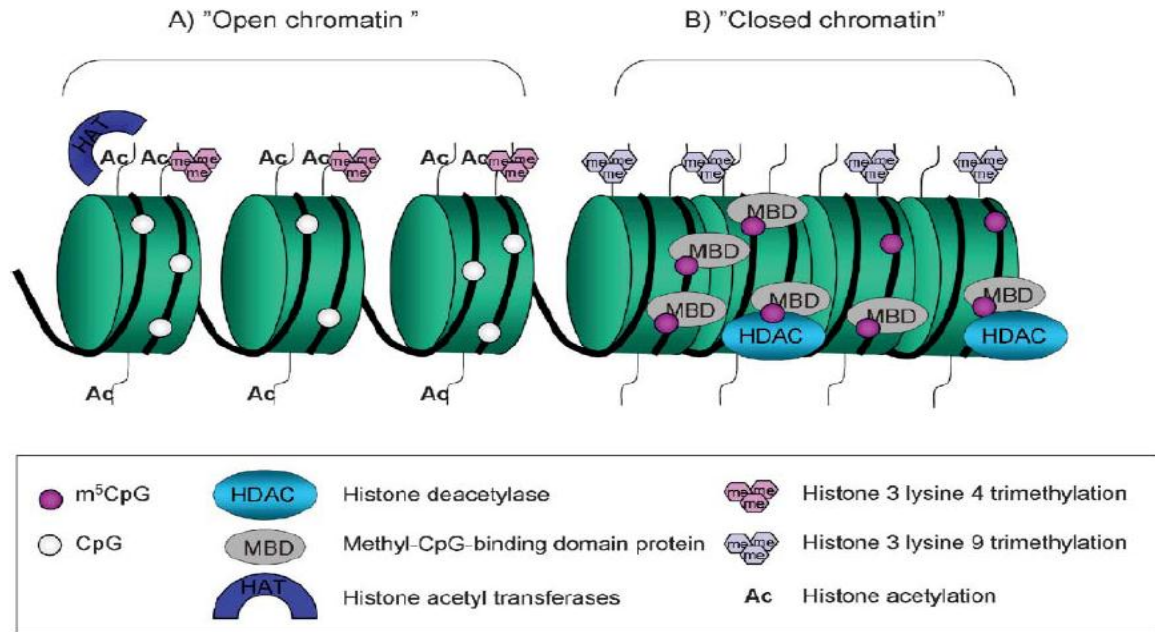


Figure 1.7. Chromatin structures of active and inactive promoters. A. Actively transcribed open chromatin structure. The CpG sites in the promoter regions are un-methylated, histone tails are acetylated and the lysine 4 on histone H3 is trimethylated. B. Transcriptionally suppressed closed chromatin regions. The promoter regions are methylated, acetylation of the histone tail is removed via histone deacetylases and replaced by two other markers of epigenetically suppressed DNA, histone H3 lysine 9 trimethylation and methyl-CpG-binding-domain proteins (MBD) (Gronbaek et al., 2007).

Type of histone modification	Symbol	Epigenetic marker
Trimethylation of lysine 9 of N-terminal tail of H3	H3K9me3	Suppressed
Trimethylation of lysine 27 of N-terminal tail of H3	H3K27me3	Suppressed
Trimethylation of lysine 36 of N-terminal tail of H3	H3K36me3	Suppressed
Trimethylation of lysine 20 of N-terminal tail of H4	H4K20m3	Suppressed
Trimethylation of lysine 4 of N-terminal tail of H3	H3K4me3	Active
Trimethylation of lysine 79 of N-terminal tail of H3	H3K79me3	Active

Table 1.2. The methylation of the N-tails of histones is used as an epigenetic signature for predicting the transcription status of the genes.

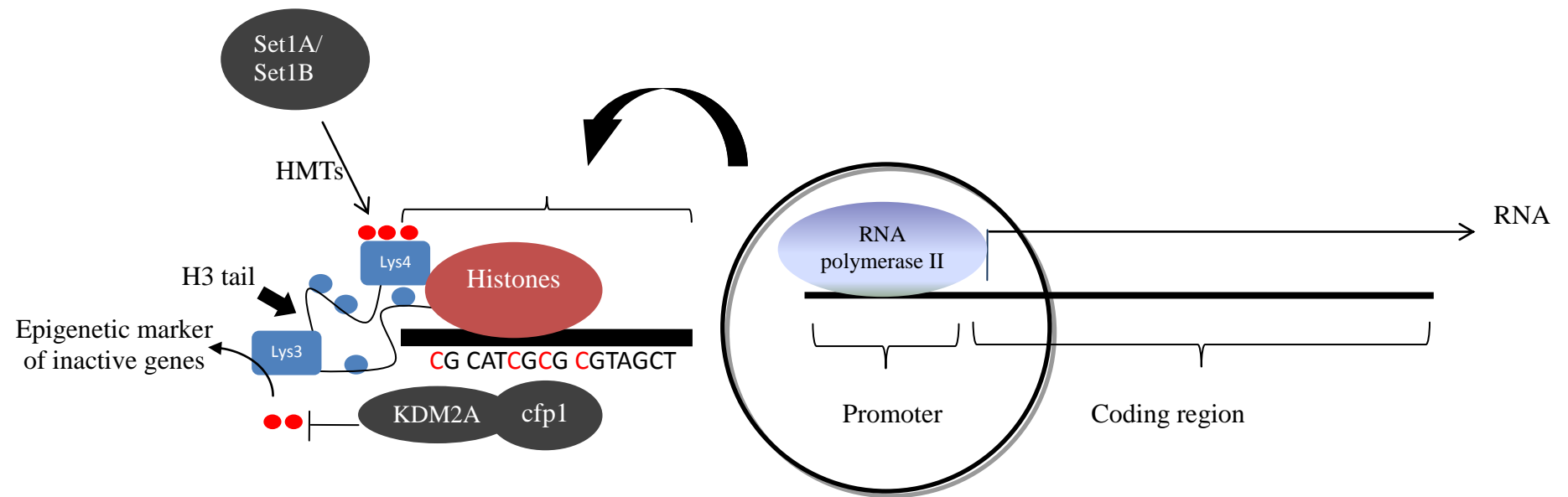


Figure 1.8. A schematic representation of the key DNA methylation and histone modifications that occur in the promoter region of a transcriptionally active gene. KDM2A is a H3K36 demethyltransferase. **C**: un-methylated cytosine, red circle: methyl group, blue circle: acetyl group, Lys: lysine, HMT: histone methyltransferase.

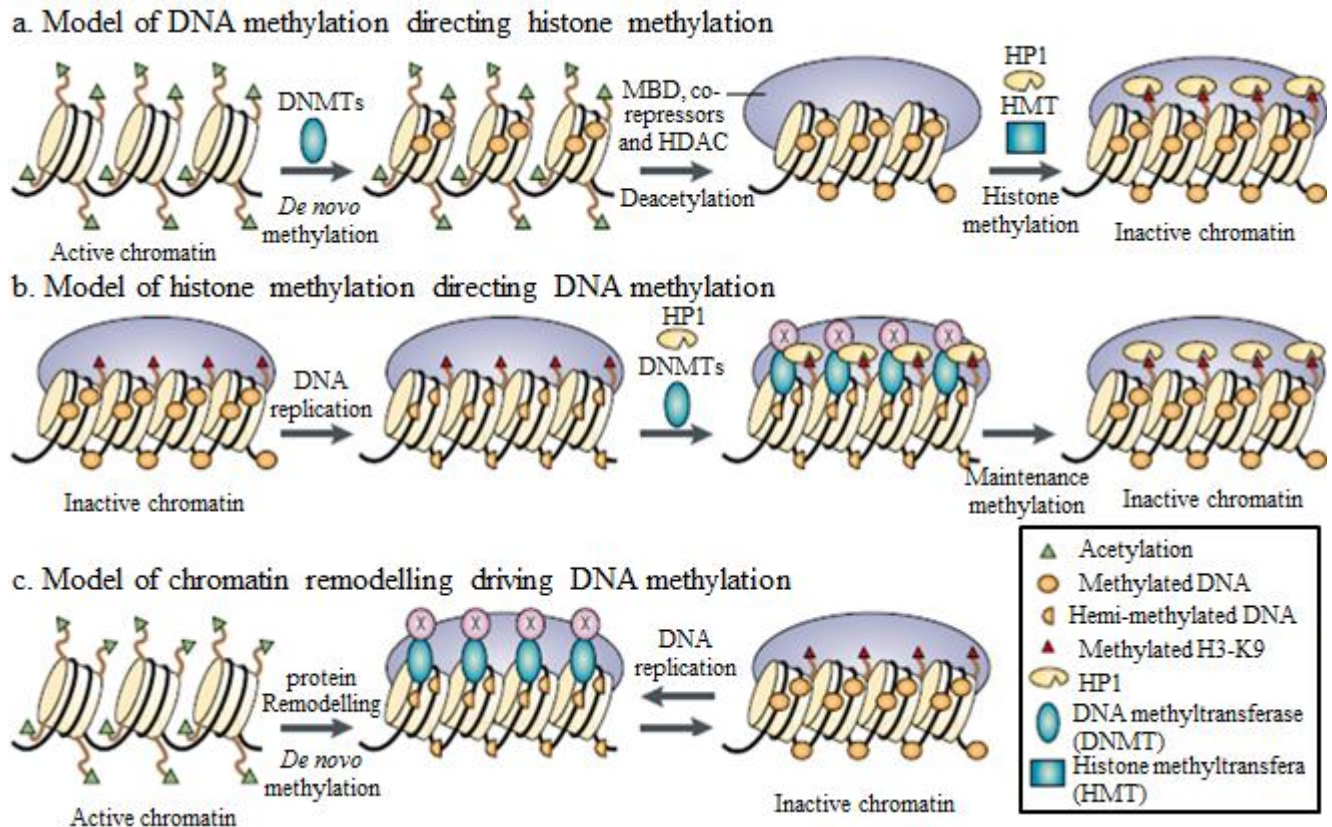


Figure 1.9. The link between DNA methylation, histone modifications and chromatin remodelling. DNA methylation and histone modifications are interdependent events regulating transcription. Three models have been proposed to explain the possible order of the events. Model 1. Model of DNA methylation directing histone methylation. DNA is methylated by *de novo* methyltransferases (DNMT3a and DNMT3b) and the DNA methylation pattern is conserved throughout replication by the maintenance methyltransferase, DNMT1. Binding of MBD-containing proteins, such as MeCP2, result in recruitment of HDACs and chromatin remodelling proteins. This causes deacetylation of histones and condensation of the chromatin structure. The chromatin will further attract histone methyltransferases (HMT). This leads to methylation of epigenetic markers associated with transcriptionally silent genes such as H3-K9. Model 2. Model of histone methylation directing DNA methylation. In this model methylation of histone tails via HMT results in recruitment of proteins such as HP1. Binding of HP1 to methylated H3-K9 regions will further direct DNMTs to these regions, resulting in methylation of DNA. Model 3. Model of chromatin remodelling driving DNA methylation. Remodelling of chromatin via ATP-dependent remodelling complexes such as Lsh can result in unwinding of DNA. This provides access for histone modification enzymes (HMT and HDAC) and DNMTs to these regions, leading to transcriptional alterations. This figure was modified from Li, 2002.

1.3.2.3. General inverse correlation between gene expression and DNA methylation

The relationship between gene expression levels and DNA methylation is extremely complex and site sensitive. When determining the influence of DNA methylation upon transcription, the location of DNA methylation in respect to the transcription start site and other regulatory regions (i.e. enhancers and suppressors) is very important (Siegfried and Simon, 2010; Jones, 1999; Van Vlodrop, 2011; Jones and Takai, 2001). For example, as discussed in section 1.3.2.1, although the region prior to the TSS of the *IGF2* gene is methylated, this gene is transcriptionally active. The methylation of this region inhibits the binding of the CTCF insulator-binding protein; thereby allowing access to, and binding of the corresponding enhancers and activation of transcription (Hark et al., 2000).

Furthermore, it has become apparent that the methylation levels of specific CpG dinucleotides, known as core CpG sites, within the CGIs of a promoter region of a gene are key factors in deciding the outcome of DNA methylation upon gene expression. Therefore, the methylation levels of these core CpG sites, irrespective of the levels and the direction of methylation (hypo- or hyper-methylated) in the remaining CpG sites surrounding the core CpG regions, will dictate and regulate transcription. In some cases core CpG sites have been narrowed down to a single CpG dinucleotide within a CpG island. This shows that regulation of transcription does not require change in methylation of an entire CGI. Indeed, methylation of a few key and gene-specific CpG sites can efficiently regulate transcription (Van Vlodrop, 2011). For example, the *tristetraprolin (TTP)* gene acts as a tumour suppressor gene by destabilising c-myc protein and inducing apoptosis in cells with DNA damage and it is commonly suppressed in liver cancers. Recently, transcriptional suppression of this gene has been linked to hypermethylation of a single CpG located at the promoter region of this gene. Methylation of this particular CpG site allows binding of a transcription suppressor complex to the promoter region of this gene and therefore prevents transcription (Sohn et al., 2010).

The presence of core CpG sites within alternative promoters at intra- and intergenic CpG dinucleotides, the regulation of transcription of several genes by one promoter and the location of several core CpG sites outside the traditionally defined CGIs can further complicate the relationship between regulation of gene expression and DNA methylation levels (Van Vlodrop, 2011). Moreover, the level of methylation required for inactivation of each gene varies (not a simple case of all or null methylation). Furthermore, these effects can

be tissue, cell and cancer type specific. This adds to the complexity and the difficulty of interpreting the DNA methylation data in regards to their effect on gene expression (Van Vlodrop, 2011).

1.3.3. The one-carbon cycle

S-adenosylmethionine (SAM) as a universal methyl donor (Pogribny, 2010) is the final and immediate source of methyl groups for methylation of a range of substrates including DNA, RNA, xenobiotics, phospholipids, hormones and neurotransmitters (Struys et al., 2000) and therefore has a vital role in the regulation of crucial biological functions (Krijt et al., 2009). SAM is produced as a result of a series of reactions that comprise the one-carbon cycle (Herceg and Vaissiere, 2011). A major function of SAM is providing the methyl group necessary for methylation of DNA which is catalysed by DNA methyltransferases. In return, SAM is converted to *S*-adenosylhomocysteine (SAH) (Kumar et al., 2008). SAH, due to its high binding affinity for the catalytic regions of SAM-dependent methyltransferases, is a potent inhibitor of transmethylation reactions (Lu, 2000; Lieber and Packer, 2002; James et al., 2002). Therefore, constant removal of SAH from the site of production is necessary for continuation of methylation reactions. Removal of SAH is achieved in a reversible reaction catalysed by *S*-adenosylhomocysteine hydrolase (SAHH) protein which results in the production of homocysteine (Melnik et al., 2000; Lu, 2000; Lieber and Packer, 2002) (An overview of the one-carbon cycle is presented in Figure 1.10). Changes in concentrations of SAH and SAM, result in disruption of SAM/SAH ratio. This ratio is referred to as the methylation index. SAH-dependent decrease in the methylation index interferes with and inhibits methylation reactions, including DNA methylation (Herceg and Vaissiere, 2011; Krijt et al., 2009; Melnik et al., 2000).

The produced homocysteine can either be converted to cysteine through a transsulfuration reaction or can be converted back to methionine through addition of a methyl group. The latter is catalysed by two independent reactions: 1. Folate, vitamin B12-dependent methionine synthesis. 2. Betaine-homocysteine methyltransferase (BHMT) reaction. In the first pathway the primary source of the methyl group is provided by dietary intake of folate. Folate is converted to 5-methyl tetrahydrofolate (5-methyl THF) which subsequently donates a methyl group to homocysteine. This results in production of methionine and tetrahydrofolate

(Emmert et al., 1996; Herceg and Vaissiere, 2011). In addition, conversion of homocysteine to methionine is achieved by BHMT. This results in the transfer of a methyl group from betaine, a product of choline oxidation, to homocysteine and generation of dimethylglycine and methionine. In contrast to methionine synthesis that occurs in all cells, expression of BHMT is mainly restricted to liver tissues in mammals. Nevertheless, both pathways contribute equally to regeneration of methionine from homocysteine in the human liver (Emmert et al., 1996). Furthermore, conversion of methionine to SAM is catalysed by methionine adenosyltransferases (MAT) in the presence of ATP.

Disruption in the one-carbon cycle by, for example diets deficient in folate and choline, the two primary sources of methyl groups for generation of SAM, result in decreased levels of SAM. This causes DNA hypomethylation in mouse livers with a subsequent consequence of liver tumourigenesis (influence of diet, change in DNA methylation and change in phenotype is further discussed in section 1.3.5.2) (Herceg and Vaissiere, 2011; Ghoshal and Farber, 1984).

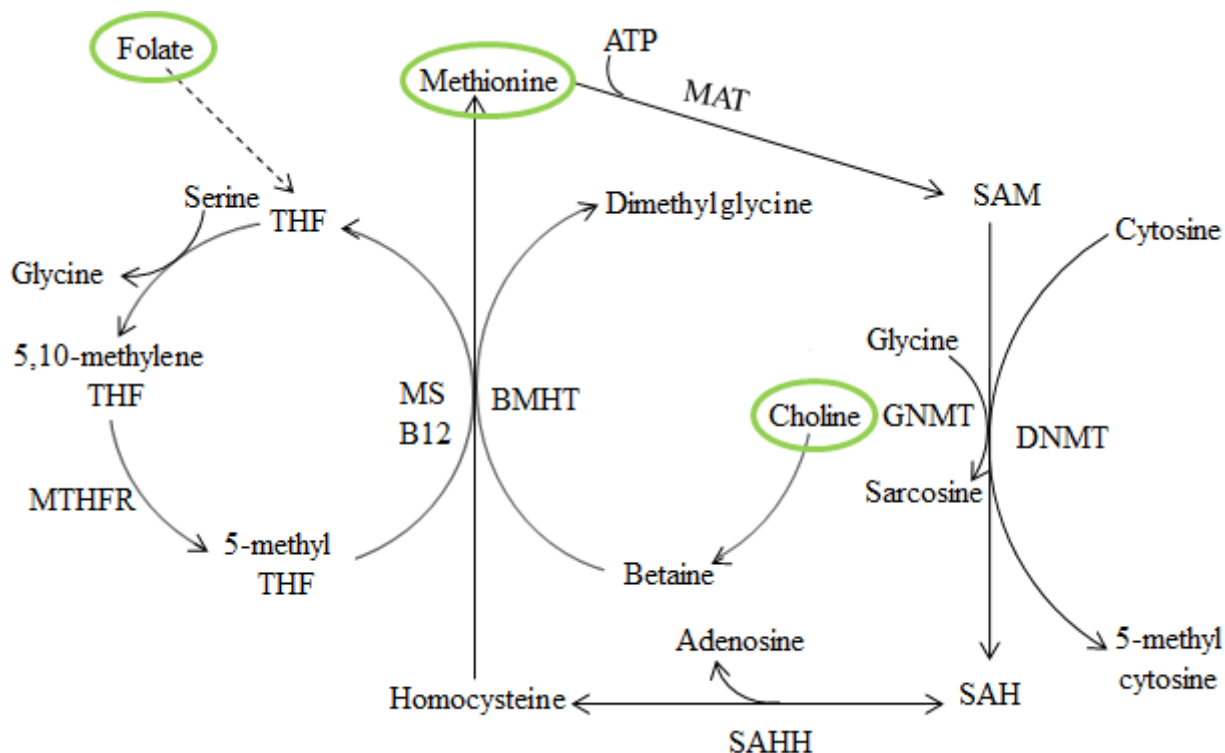


Figure 1.10. The one-carbon cycle. In this pathway ATP activates methionine which results in the generation of the universal methyl donor, SAM. SAM is used in a range of biological reactions including methylation of DNA. The one-carbon cycle is highly regulated. Changes in components of this pathway can cause decrease in the SAM/SAH ratio which, subsequently leads to DNA hypomethylation. For example, reduced levels of essential nutrients, such as folate and choline, and accumulation of homocysteine lead to a decreased SAM/SAH ratio. Furthermore, sufficient levels of SAM trigger mechanisms that lead to conversion of SAM to SAH. For example, 5-methyl THF can inhibit GNMT activity and SAM can regulate GNMT activity by inhibiting the synthesis of 5-methyl THF. Therefore, sufficient levels of SAM result in SAM-dependent inhibition of 5-methyl THF synthesis which subsequently results in GNMT-dependent SAM to SAH conversion and reduction of SAM/SAH ratio.

Abbreviations: SAM: S-adenosylmethionine, SAH: S-adenosylhomocysteine, DNMT: DNA methyltransferase, GNMT: glycine N-methyltransferase, SAHH: S-adenosylhomocysteine hydrolase, THF: tetrahydrofolate, MS: methionine synthase, B12: vitamin B12, BHMT: betaine-homocysteine methyltransferase, MTHFR: methylenetetrahydrofolate reductase, MAT: methionine adenosyltransferases. Green circle: essential nutrients which are mainly obtained through diet (Ulrey et al., 2005; Melnyk et al., 2000; Lu, 2000; Herceg and Vaissiere, 2011).

1.3.4. Epigenetic regulation of normal functions in the cells

As DNA methylation patterns are maintained post-replication, these modifications contribute to the regulation of transcription of normal biological functions of cells such as inactivation of the X-chromosome, imprinting, suppression of transcription of parasite sequences, differentiation and tissue specific gene expression (Jones and Takai, 2001) which are discussed briefly below. In addition, as DNA methylation influences protein-DNA interactions and as many pathways are regulated through DNA-protein interactions, the vital role of DNA methylation in many biological functions is apparent (Doerfler, 1983).

1.3.4.1. Imprinting

DNA methylation is involved in imprinting (Gronbaek et al., 2007). Imprinting or parent-of-origin specific-gene expression is the expression of one of the two alleles of a gene, either the maternal or the paternal allele, and complete suppression of the other allele. This causes non-equivalence between mammalian parents in autosomal genetic material and prevents parthenogenesis (Reik and Walter, 2001; Hore et al., 2007). Imprinting has only been confirmed in mammals, with approximately 100 imprinted genes identified in humans and mice. Imprinted genes are usually located in clusters, partly as a result of a phenomenon known as epigenetic spreading (Reik and Walter, 2001).

In addition to preventing parthenogenesis and possible correction of gene duplications, it is hypothesised that imprinting mechanisms have evolved in mammals to prevent conflict between offspring and mother over food resources. This is referred to as the “parent conflict hypothesis”. Some of the paternal genes increase foetal growth at the expense of the mother, whilst it is in the interest of the mother to limit this growth. Therefore, imprinting is a way of balancing foetal growth rate. For example in mice *IGF2*, a growth enhancer, is imprinted in the maternal allele while the *IGF2* receptor, a growth suppressor, is methylated in the paternal allele. However, due to the spreading of the epigenetic changes some of the imprinted genes are “innocent bystanders” in this conflict (Hore et al., 2007).

As mentioned in section 1.3.1, mammalian genomes undergo two extensive epigenetic reprogramming events, one during embryogenesis and the other during gametogenesis. In the latter epigenetic markers, including DNA methylation and chromatin modifications are re-set

in a sex-specific manner in imprinted genes which are protected from the second wave of reprogramming during embryogenesis (Hore et al., 2007; Godmann et al., 2009). This results in inheritance of imprinted patterns in a parent-of-origin manner in the somatic cells (Reik and Walter, 2001). Therefore, as imprinted genes are only represented and expressed monoallelically, methylation or deletion of the active allele has serious consequences. For example Prader-Willi and Angelman syndrome are both consequence of deletion of the active paternal and maternal alleles in the 15q11-13 region of the neuronal tissue, respectively (Hore et al., 2007).

1.3.4.2. X-chromosome inactivation

DNA methylation is involved in X chromosome inactivation (Gronbaek et al., 2007). During this process, in order to compensate for the 2-fold higher occurrence of the genes located on the X chromosomes in females compared to males, one of the X chromosomes in females is inactivated (X_i) (Li, 2002). The X inactivation occurs during embryogenesis after implantation, with equal inactivation probabilities for the maternal and paternal X chromosome (Avner and Heard, 2001; Li, 2002).

The inactivation process is initiated by X (inactive)-specific gene (*Xist*) located at the X inactivation centre (XIC). The *Xist* gene encodes a large non-coding RNA, expressed at a background level from both alleles prior to X chromosome inactivation. During inactivation a marked increase in expression and accumulation of *Xist* gene of the allele selected for inactivation is observed. This leads to cis coating of the X chromosome with the *Xist* RNA and triggers a cascade of events leading to inactivation of the entire X-chromosome (Avner and Heard, 2001).

X chromosome inactivation is a synergy between DNA methylation, histone deacetylation, chromatin remodeling and *Xist* gene expression. Methylation of the *Xist* gene of the active X chromosome (X_a) prevents its inactivation (Norris et al., 1994). In addition, following initiation of the inactivation via *Xist*, DNA methylation contributes to the inactivation process by spreading the inactivation to all CGIs of the housekeeping genes in the X chromosome and stabilising and ensuring maintenance of a silence phenotype throughout cell division (Avner and Heard, 2001). The significance of the contribution of other epigenetic mechanisms is extremely clear. This is because the X_i resembles a highly condensed heterochromatin

comprised of CGI regions containing methylated CpG dinucleotides and deacetylated histones in the promoters of the house keeping genes (Csankovszki et al., 2001). In addition, it has been shown that mutation in *DNMT3b* gene, without altering the activity of Xist gene, results in hypomethylation and activation of several X-linked genes (Li, 2002).

1.3.4.3. Differentiation and tissue specific gene expression

The DNA sequence of the genetic material is the same for all tissue types, thus it cannot directly explain the differences observed in expression patterns and phenotypes of diverse categories of tissues. Phenotype-specific gene expression in differentiated cells is achieved mainly by altering the proportion and location of hetero- and eu-chromatin regions. As epigenetic markers regulate gene expression by altering the protein-DNA interactions and DNA packaging, epigenetic mechanisms have been proposed as the key components for establishing tissue-specific gene expression profiles (Godmann et al., 2009; Enrlich, 2005; Momparler and Bovenzi, 2000). In addition, epigenetic markers in contrast to gene sequences are flexible. Therefore, the epigenomes of different cells can differ depending on the microenvironment of differentiating cells (Godmann et al., 2009). Indeed, tissue-specific expression of genes with altered methylation profiles (hypo- or hyper-methylation) based on the tissue type, have been identified in humans (Christensen et al., 2009). For example, testes and uterus exclusively express a testis-specific lactate dehydrogenase gene and a myometrium-specific oxytocin receptor respectively and these are controlled through methylation of their regulatory regions (Enrlich, 2005). Although, it has to be stated that, due to the complex nature of the relationship between DNA methylation and gene expression (discussed in section 1.3.2.3), the contribution of epigenetic mechanisms is not always clear.

1.3.4.4. Transposons

Transposons, scattered throughout the genome, comprise more than 40% of the mammalian genome (compared to exonic regions comprising <2%) (Jones and Takai, 2001). Due to evolution-induced alterations, most transposable elements are inactive. However, the remaining small proportions of active transposons are a threat to the genome by inducing mutations, chromosome instability, translocations and formation of chimeric mRNAs. Nevertheless, generally, little damage is caused by these potentially hazardous elements. This

has been mainly attributed to DNA methylation-induced inactivation of transposons. This theory, known as the genome defence model, corresponds with the CpG dinucleotide richness of the promoter regions of transposons (Yoder et al., 1997; Gronbaek et al., 2007; Esteller, 2008). Indeed, the mammalian genome is found to be globally methylated (Suzuki and Bird, 2008) with the most highly methylated regions corresponding to transposons (Bird, 2002).

In contrast to high levels of methylated transposable elements in somatic cells, transposons are un-methylated and transcriptionally active in germ cells, where they can cause long-term damage compared with somatic cells (Bird, 2002). Therefore, in addition to the genome defence model proposed by Yoder et al., (1997), it has been suggested that in general, methylation of un-necessary promoters may have evolved as a mechanism to reduce the background transcriptional noise (Bird, 2002).

1.3.5. Cancer as an epigenetic disease

Epigenetic mechanisms are implicated at all stages of tumourigenesis, from initiation to metastasis. There is growing certainty that epigenetic mechanisms, including DNA methylation, precede genetic changes (Sharma et al., 2010; Herceg and Vaissiere, 2011; Moggs et al., 2004; Feinberg et al., 2006; Pogribny, 2010). Indeed, alterations in the epigenome trigger tumourigenesis and result in stimulation of a cascade of genetic and additional epigenetic changes. Therefore, cancer is as much an epigenetic disease as it is a genetic disease (Herceg and Vaissiere, 2011).

1.3.5.1. DNA methylation and cancer

Distortion of the DNA methylation profile is a key event in tumourigenesis and a known hallmark of tumour cells. Detection of abnormal low levels of global methylation in tumours was one of the first identified epigenetic events in human cancers. This revealed loss of methylation at repetitive DNA sequences, coding and intronic regions (Esteller, 2008). Indeed, both global DNA hypomethylation and gene-specific hypermethylation (e.g. genes involved in apoptosis, metastasis, adhesion, cell cycle and DNA repair) are observed in all investigated human cancers (Lopez et al., 2009; Tischoff and Tannapfel, 2008; Gronbaek et al., 2007; Esteller and Herman, 2002; Esteller et al., 2001).

It was thought that epigenetic alterations, such as change in DNA methylation of tumours are consequences of altered gene expression in tumours (Feinberg, 2001). However, it is now known that alterations in epigenome, which can be induced by several factors such as chemicals, can initiate changes in gene expression with consequences such as formation of tumours. Formation of tumours can further trigger change in the epigenome (Pogribny, 2010). Therefore, it should be considered that some epigenetic changes initiate the neoplastic transformation (i.e. they are initiated by the agent) and some are characteristic of the formed tumour. DNA methylation contributes to the tumourigenesis process through several mechanisms as shown in Figure 1.11 (Herceg and Vaissiere, 2011; Esteller and Herman, 2002; Esteller, 2008; Herrera et al., 2008).

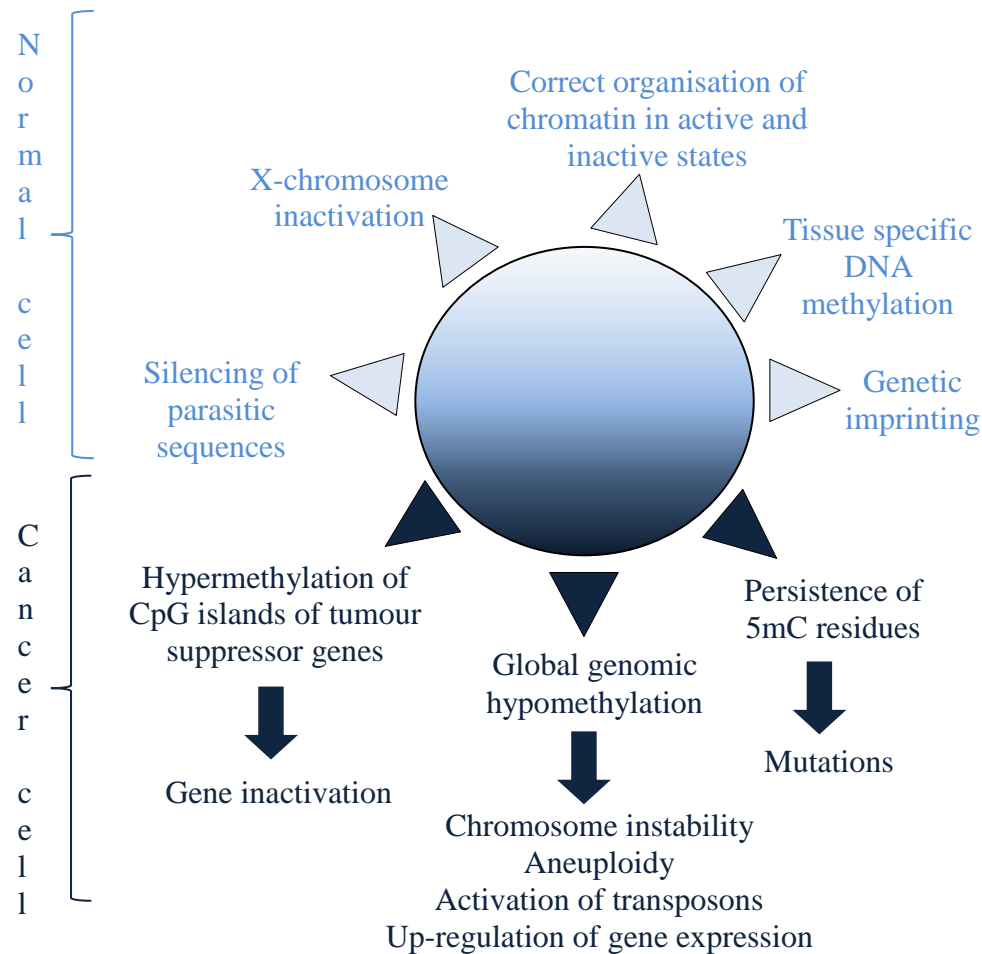


Figure 1.11. Biological significance of DNA methylation in normal and tumour cells. This figure was modified from Esteller and Herman, 2002.

1.3.5.1.1. DNA hypomethylation

Large scale de-methylation of the genome was one of the first identified epigenetic phenomena in 1983 implicated in tumourigenesis (Gama-Sosa et al., 1983). Since then, it has been recognised as a common feature in all tumours (Feinberg et al., 2006). DNA hypomethylation promotes tumour development through a number of methods including: 1) Activation of oncogenes and genes associated with metastasis and tumour aggression, 2) Activation of repetitive sequences, 3) Chromosomal instability, 4) Aneuploidy and recombination and 5) Biallelic transcription of imprinted genes (Herceg and Vaissiere, 2011).

DNA methylation levels are reduced 20%-60% in cancer cells compared to normal cells. This mainly reflects the DNA hypomethylation detected at coding regions and introns of genes and the repetitive regions (Esteller and Herman, 2002). Considering that transcription of repetitive regions is primarily suppressed through their heavy methylation and the fact that approximately half of the human genome is comprised of repetitive regions, including mobile transposons and satellites that are dispersed throughout the entire genome, it is not surprising that most of the observed DNA hypomethylation in cancers corresponds to these regions (Ehrlich, 2009; Herceg and Vaissiere, 2011). Thus hypomethylation of transposable elements and pericentromeric repeats results in their activation, followed by instability and rearrangement of the chromosomes and aneuploidy (Esteller and Herman, 2002; Stratthdee and Brown, 2002; Ehrlich, 2009).

Other than hypomethylation of the whole genome, hypomethylation of specific genes such as oncogenes, genes associated with metastasis, and drug resistance genes have been detected in mammalian tumours (Ehrlich, 2009). For example the *BCL-2* oncogene and the gene encoding urokinase-type plasminogen activator (*uPA*) are hypomethylated and over-transcribed in human B-cell chronic lymphocytic leukaemia and in breast cancer with increased aggressiveness, respectively (Laird and Jaenisch, 1996; Ehrlich, 2009).

Another consequence of DNA hypomethylation in tumours is loss of imprinting (LOI). LOI results in biallelic expression of the previously imprinted genes, hence increasing the expression levels of these genes by two-folds (Laird and Jaenisch, 1996). As most of the imprinted genes are related to biological functions such as regulation of cell growth, cell signalling, cell cycle and transport; their over-expression has major consequences in normal development and function of a cell and has been implicated in many human disorders, such as cancers (Feinberg et al., 2006). For example LOI of *IGF2* gene is correlated with five-fold increase in the risk of colorectal cancer and is commonly linked with Wilms' tumour (Herceg and Vaissiere, 2011; Laird and Jaenisch, 1996).

The mechanism of DNA hypomethylation has been largely unexplored. This is due to the nature of hypomethylated regions and their high repetitiveness. However, recent emergence of whole genome methylation profiling techniques and discovery of single copy-hypomethylated genes in tumours, known as germ cell-specific genes or cancer-germline (CG) genes, has led to recognition of DNA hypomethylation in tumours as a selective event rather than a non-

selective, random procedure (De Smet and Lorient, 2010; De Smet et al., 2004). Thus regions between 1kb to several megabases undergo hypomethylation in cancers while other regions escape these waves of DNA hypomethylation. This leads to formation of a mosaic DNA methylation pattern (De Smet and Lorient, 2010). As a result, DNA hypomethylation patterns are tumour-type specific. In fact, activation levels, types of CG genes and some repetitive sequences have shown tumour specificity (De Smet and Lorient, 2010; Akers et al., 2010). For example the *NBL2*, a DNA repeat in the acrocentric chromosomes, is hypomethylated in neuroblastomas and hepatocellular carcinoma and is hypermethylated in ovarian cancer and Wilms' tumours while showing intermediate methylation levels in normal tissues (De Smet and Lorient, 2010).

So far about 50 CG genes have been identified in humans (De Smet and Lorient, 2010). As these genes are hypermethylated in normal somatic cells and hypomethylated in germ cells, their expression is restricted to germ cells. Although the functions of all these genes are not entirely clear, it is thought that they express tumour-specific antigens and may be involved in providing them with stem cell-like properties such as immortality (De Smet and Lorient, 2010; Akers et al., 2010; Lorient et al., 2006). In addition, there has been some evidence that CG antigens are hallmarks of normal stem cells. This supports the stem cell origin of tumours, as tumour cells also express these antigens (this model is explained in section 1.3.5.3). However, the role of CG genes in stem cells and provision of the tumours with stem cell-like properties is inconclusive and requires further investigation (Akers et al., 2010).

Several models for mechanisms of DNA hypomethylation in tumours have been proposed. One model that seems most likely is the demethylation/remethylation model. This model is based on the same principals observed in reprogramming events during embryogenesis. The entire genome of a tumour cell undergoes non-selective, active demethylation. This is followed by a selective *de novo* methylation of the genome. This re-methylation only takes place at regions that are not protected by transcription factors. Therefore, the selectiveness of DNA hypomethylation is based on a protection mechanism against re-methylation (Figure 1.12) (De Smet and Lorient, 2010). For example, the melanoma antigen-A1 (MAGE-A1) CG gene is methylated and in-active in somatic cells while it is hypomethylated and active in tumours. Therefore, MAGE-A1 becomes de-methylated and it is protected against re-methylation by *de novo* methyltransferases in tumour cells. This protection against DNA

methylation is achieved by binding of E-twenty six (ETS)-TF to the regulatory regions of this gene, thus inhibiting the access of *de novo* methyltransferases to these regions. It has been shown that removal of ETS transcription factor results in methylation of MAGE-A1 in tumour cells (De Smet et al., 2004; De Smet and Lorient, 2010).

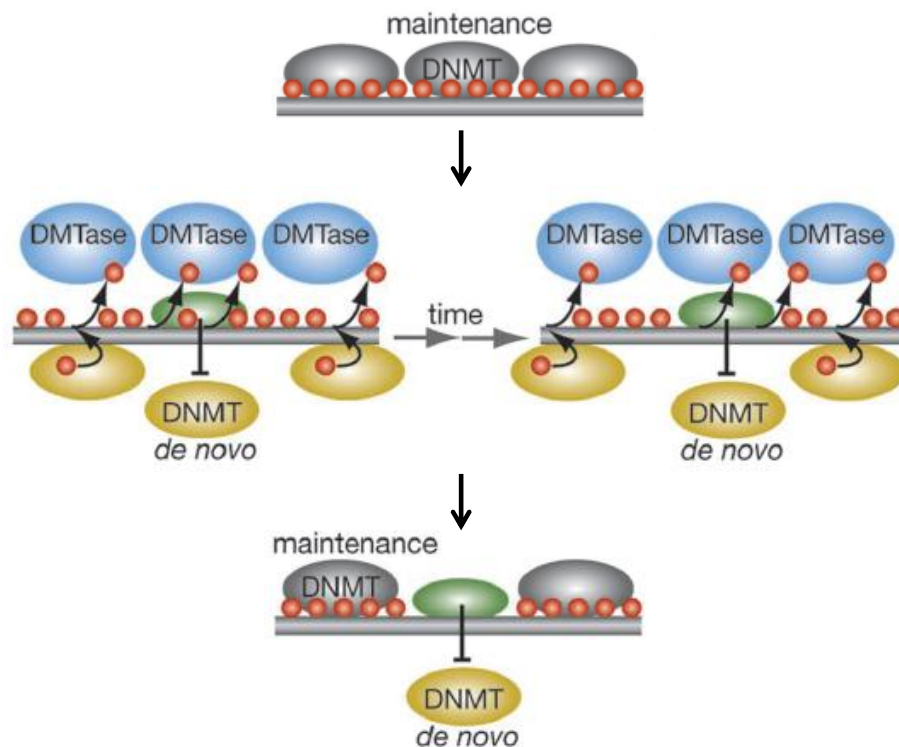


Figure 1.12. Demethylation/remethylation model for site specific DNA hypomethylation in tumours. Methylation is removed non-selectively from the entire genome. Methylation is restored by *de novo* methyltransferases at the regions that are not protected from re-methylation by transcription factors. This leads to selective hypomethylation. This model is one of the proposed models for DNA hypomethylation in tumours. However, based on current evidence, it appears that it is also the most likely model. Red circles: methylated cytosine, green oval: transcription factor, blue oval: active demethylase (the nature of these enzymes is not known), gray oval; maintenance methyltransferase, yellow oval: *de novo* methyltransferase. This image was modified from De Smet and Lorient, 2010.

1.3.5.1.2. DNA hypermethylation and inactivation of tumour suppressor genes

Methylation of the core CpG dinucleotides located near the transcription start sites of tumour suppressor genes (TSG), which are not normally methylated, results in their inactivation. The significance of the influence of TSG inactivation in all pathways of tumourigenesis is well established (Jones and Takai, 2001; Gronbaek et al., 2007).

Although the mechanism of suppression of transcription by DNA methylation is becoming more evident, the factors that initiate the *de novo* methylation of TSGs in cancers and the preferential methylation of groups of genes based on tumour types, are still poorly understood (Herceg and Vaissiere, 2011; Esteller and Herman, 2002). Several hypotheses have been proposed, such as over-expression of *DNMTs* and defects in the mechanisms responsible for the protection of CpG dinucleotides against methylation in somatic cells or in stem cells (Laird and Jaenisch, 1994; Herceg and Vaissiere, 2011). As discussed, this change in methylation of TSGs can arise as early as the differentiation stage in somatic stem cells. It is well established that disruption of methylation during stem cell differentiation (epigenetic mechanisms are key in this process) can result in abnormal methylation of genes. This results in generation of a population of pre-neoplastic progenitor cells with growth and survival advantages compared to other cells (Laird and Jaenisch, 1994). This is part of the epigenetic progenitor model of tumourigenesis which is described in section 1.3.5.3.

The list of hypermethylated genes in tumours is continuously growing. Some of the hypermethylated genes that are involved in crucial pathways and are known to be altered in cancers are shown in Table 1.3.

Biological function	Gene
Cell cycle and signalling	<i>P16/CDKN2A</i> (most tumours), <i>P15/CDKN2B</i> (leukaemia, gastric cancer, hepatocellular tumours), <i>RB</i> (retinoblastoma), <i>APC</i> (Breast, lung, colon, gastric, oesophageal, pancreatic and hepatocellular tumours)
DNA repair	<i>MGMT</i> (colon, lung, brain, oesophageal and gastric cancer), <i>MLH1</i> (gastric, colon, ovarian and endometrial cancer), <i>BRCA1</i> (breast and ovarian cancer)
Apoptosis and survival	<i>DAPK1</i> (lymphoma and lung cancer)
Detoxification	<i>GSTP1</i> (prostate, breast and kidney cancer)
Cellular adhesion	<i>E-cadherin</i> (leukaemia, breast, lung, colon, gastric and prostate cancer), <i>H-cadherin</i> (ranges of tumours, including breast cancer)
Transcription regulation	<i>GATA4</i> , <i>GATA5</i>

Table 1.3. Examples of identified hypermethylated genes in human cancers associated with key pathways in tumourigenesis (Herceg and Vaissiere, 2011; Laird and Jaenisch, 1994; Esteller and Herman, 2002; Strathdee and Brown, 2002). In the cases where hypermethylation of a gene has been highly associated with a specific tumour type, this has been indicated. Abbreviations; *CDKN2A*: cyclin-dependent kinase inhibitor 2A; *CDKN2B*: cyclin-dependent kinase inhibitor 2B; *RB*: retinoblastoma; *APC*: adenomatous polyposis coli; *MGMT*: O-6-methylguanine-DNA methyltransferase; *MLH1*: MutL homolog 1; *BRCA1*: breast cancer 1, early onset; *DAPK1*: death-associated protein kinase 1; *GSTP1*: glutathione S-transferase P1; *GATA4*: GATA binding protein 4; *GATA5*: GATA binding protein 5.

1.3.5.1.3. Increased rate of mutations

As demonstrated for the first time in the genome of *Escherichia coli* by Coulondre et al., (1978), endogenous and exogenous rates of mutations are increased at 5-methylcytosine sites (Weisenberger and Romano, 1999). Deamination occurs at the 4th carbon of both cytosine and 5-methylcytosine (5-mC). It has been estimated that 100 to 500 cytosine deamination reactions occur each day within each cell (De Bont and Van Larebeke, 2004). However, the rate of 5-mC deamination compared to cytosine is 3 to 4 times higher. Deamination of cytosine results in an unusual base composition of DNA (C to U transition); hence, it is more readily recognised and rapidly repaired by the abundant uracil-DNA glycosylase enzyme. This enzyme removes the uracil base, resulting in generation of a base-free site which is

further repaired. However, 5-mC deamination results in 5mC to T transition which is recognised and repaired less effectively and with a slower rate than C to U transition by the mismatch repair enzymes (Gronbaek et al., 2007; De Bont and Van Larebeke, 2004). Indeed, the higher instability of 5-mC and less effective repair of 5mC:T transitions compared to C:U transitions is reflected in the genome base composition as CpG dinucleotides are five-fold under-represented in the mammalian genome compared to other dinucleotides (Laird and Jaenisch, 1994). In addition, 5mC induced point mutations account for more than 30% of disease related mutations, with genes associated with cell growth and survival comprising the majority of the identified mutated genes. For example, the cytosine bases at the three main CG:TA transition mutation sites of tumor protein 53 (*P53*), are always found to be methylated in normal tissues (Gronbaek et al., 2007).

Furthermore, 5-mC increases the binding and exogenous mutagenic effect of certain chemicals, such as esperamicins A and B and mitomycin C, at the adjacent guanine base. This increased targeting of guanine base by chemicals at 5-mC sites has been linked to structural alterations and unwinding of DNA at methylated sites. These alterations lead to increased accessibility of these regions to carcinogens (Weisenberger and Romano, 1999; Gronbaek et al., 2007).

Finally, the rate of mutations at methylated CpG dinucleotide sites are also increased through other epigenetic alterations such as DNA hypermethylation. For example, genes for several repair enzymes, such as mismatch repair proteins *MLH1* and *MSH2*, are commonly hypermethylated in colorectal cancer. Inactivation of these genes and other repair related genes, through hypermethylation or mutations, results in less effective repair of the mutations induced by 5-mC deamination; thereby, adding to the cascade of events that contribute to tumourigenesis (Gronbaek et al., 2007).

1.3.5.2. Epigenetic mechanisms and cancer: An interface between the environment and the genome

A key question is what causes the change in the phenotype of an individual as a response to its environment (i.e. chemical contaminants, diet and stress) which then can cause a higher susceptibility to certain diseases? It is becoming evident that, although epigenetic markers are stable enough to regulate gene expression, they are also susceptible to environmental signals.

This means that the epigenome can change as a response to environmental stimuli which then can lead to alteration in the phenotype (Jaenisch and Bird, 2003). In a way, the epigenome acts as a link between environmental factors (external and internal factors) and phenotype by translating the environmental signals to phenotypic responses through altered gene expression profiles. Thereby, environmental agents can affect the phenotype of an individual not only by inducing mutations and genetic variations but also by altering the epigenome. This results in changes in the expression of the genes without any alterations to DNA sequences (Dolinoy et al., 2007; Massicotte et al., 2011). Therefore, epigenetic modifications during critical stages of development (i.e. the DNA methylation re-programming stages during embryogenesis and gametogenesis) have been recognised, in part, as the factors affecting development of certain adulthood phenotypes long after the stimulating factor has been removed. These include linkage of various cancers, diabetes, obesity and behavioural and neurodegenerative disorders with environmental factors encountered during prenatal and early postnatal periods. This is known as “foetal basis or developmental origins of adult-onset of disease” (Jirtle and Skinner, 2007; Barros and Offenbacher, 2009; Jaenisch and Bird, 2003; Li et al., 2003). In a way, the epigenetic markers inflicted upon the genome by environmental factors very early on during an individual’s life act as epigenetic memories of the exposures. This epigenetic memory can manifest in adults as a pathological phenotype often following a secondary trigger such as aging or changes in hormone levels. For example, in 1950s diethylstilbestrol (DES) was used during pregnancy to prevent spontaneous abortion. However, DNA methylation changes induced by this agent during embryogenesis have been identified as the cause of a range of disorders such as increased breast and testicular cancers in adult female and male offspring (Prins, 2008). Therefore, the interactions between environmental factors and epigenetic mechanisms are extremely important in the development of diseases.

However, establishing a cause and effect between exposure to environmental factors, change in epigenome and disease is challenging. As well as the need to eliminate genetically-induced phenotypic changes in response to environmental factors, epigenetic changes as a response to environmental factors can appear to be minor. However, these changes to the epigenome can accumulate over-time leading to gradual alteration of the phenotype (Baccarelli and Bollati, 2009). For example, as shown in the viable yellow agouti (A^{vy}) mouse model, epigenetic changes on a single gene can have a significant impact on an individual’s health and on the health of following generations. Furthermore, this model is commonly used to demonstrate

the capability of environmental factors in induction of epigenetic alterations in the agouti allele at sensitive developmental stages. These changes will subsequently lead to variable expression levels of the agouti gene in genetically identical species (Dolinoy et al., 2007). In the agouti mouse model the colour of the coat is controlled by the agouti gene. The coat colour in the agouti mouse is set in their hair follicles during early stages of development. The wild-type agouti gene causes a circular production of pheomelanin (yellow pigment) and eumelanin (black-brown pigment), thus resulting in normal agouti pattern fur. Spontaneous insertion of the intracisternal A-particle (IAP) transposable element upstream of the transcription start site of the agouti gene gives rise to the A^{vy} allele. The IAP transposable element is a metastable epiallele containing a cryptic CpG rich promoter (Dolinoy, 2008; Jirtle and Skinner, 2007). Metastable epialleles are expressed at various levels in genetically identical individuals. This variation in expression level is due to epigenetic modifications that occur during early development (Dolinoy et al., 2007). The expression of the A^{vy} allele depends on the methylation levels of the IAP promoter. Maternal nutrients (i.e. methyl deficient diet) or environmental exposure at early stages of development alter the extent to which the A^{vy} allele is methylated. Unmethylated IAP promoter gives rise to an ectopic and ubiquitously expressed agouti gene from the IAP transcription start site (TSS). As the IAP-induced transcription is not restricted to the hair follicles and it is expressed throughout all cells, this gives rise to yellow fur agouti mice with high risk of adult-onset of obesity, diabetes and tumourigenesis. In contrast, complete methylation of the IAP promoter results in suppression of IAP-induced transcription, which leads to generation of pseudoagouti mice. Finally, mosaic methylation leads to formation of the mottled mice phenotype with the IAP-induced transcription in some but not all cells (Dolinoy, 2008; Jaenisch and Bird, 2003) (Figure 1.13).

The influence of environmental factors such as diet on induction of epigenetic changes at early stages of development with subsequent stable changes in gene expression and phenotype during adulthood was demonstrated using the viable yellow agouti (A^{vy}) mouse model. Dietary methyl donor supplements, such as choline, betaine and folic acid during mice pregnancy, resulted in increased methylation of the cryptic CpG rich promoter of the IAP transposable element in the offspring. Hence, the coat colour distributions of the offspring were changed towards the pseudoagouti phenotype (Jirtle and Skinner, 2007). Similar results were observed when genistein at levels comparable to the diet of a human with high soy

consumption, was fed to gestating agouti mice. This resulted in methylation of the cryptic promoter of the IAP in the developing offspring and shifting the coat colour to pseudoagouti phenotype. Moreover, it protected the offspring from adulthood obesity, diabetes and tumourigenesis (Dolinoy et al., 2006). In addition, higher proportions of A^{vy} mice are born from mothers with yellow A^{vy} phenotype. This indicates that some epigenetic markers are not entirely erased during embryonic reprogramming (transgenerational epigenetic inheritance). Hence, some epigenetic alterations induced by environmental factors (e.g. diet), not only compromise the health of the individual but they can also persist in the following generations and impact their health too (Jaenisch and Bird, 2003).

As epidemiological studies conducted in the past two decades have identified an association between the incidence rate of various cancers and certain aspects of diet (Peto, 2001; Dolinoy et al., 2007); it is plausible that, accumulation of epigenetic alterations, in part, could explain the sensitivity of tumourigenesis to environmental factors (Jaenisch and Bird, 2003). For example, folate (methionine precursor) or methionine (SAM precursor) deficient diets and induction of genome-wide DNA hypomethylation have been linked to increased risk of colorectal cancer in humans (Giovannucci et al., 1993).

Another example of the influence of environment on the epigenome and subsequent changes in gene expression is the response of *Arabidopsis* to prolonged exposure to cold weather (vernalisation). Following prolonged exposure to cold weather flowering locus C, a repressor of flowering, becomes epigenetically silenced. This results in coordination of flowering time (phenotype) with spring and summer (He et al., 2003; Bastow et al., 2004; Burn et al., 1993).

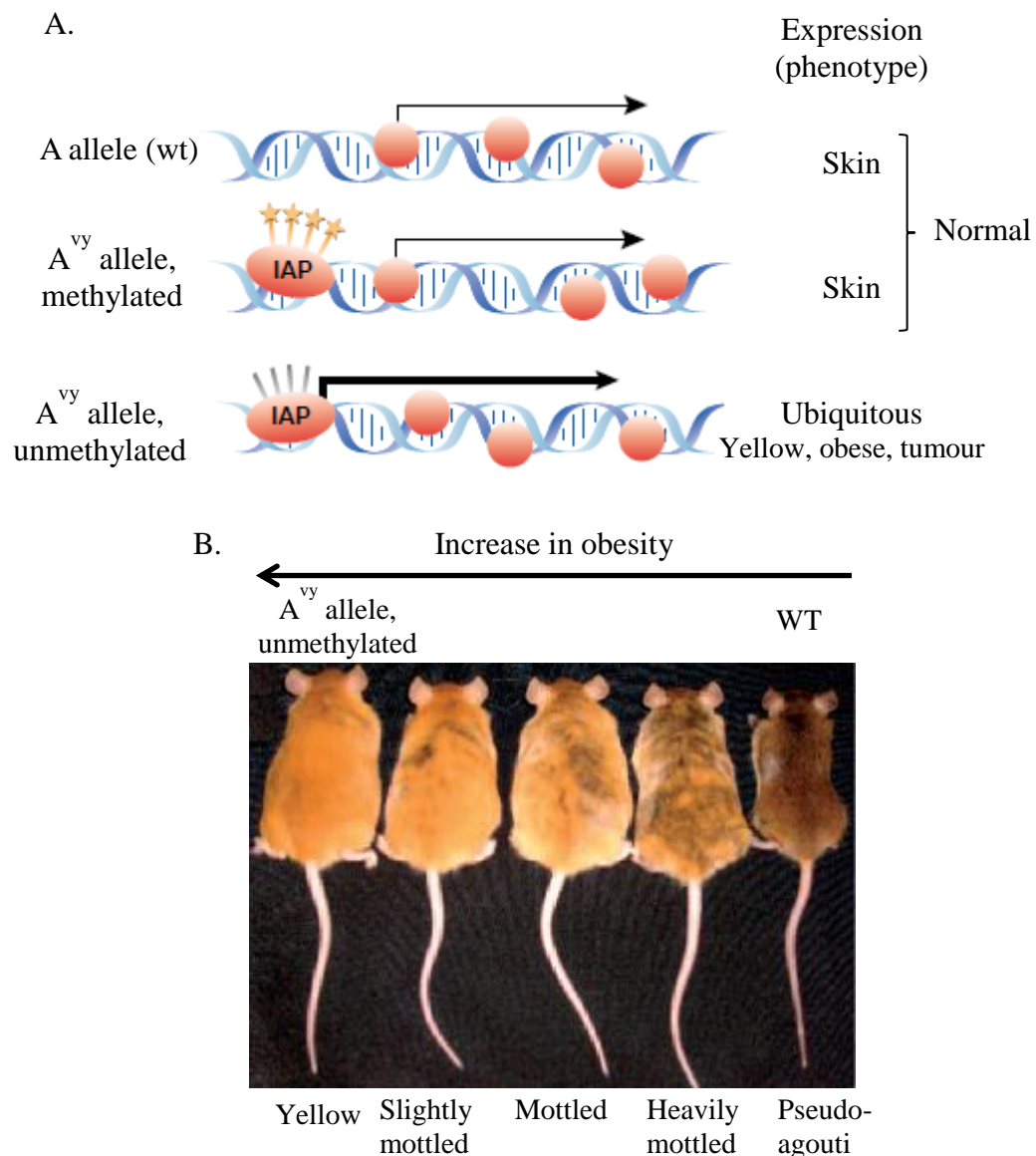


Figure 1.13. The viable yellow agouti (A^{vy}) mouse model. A. The methylation status of the cryptic CpG rich promoter of the IAP element and the associated phenotype. Methylation of the IAP promoter results in restricted expression of the agouti gene in the skin, similar to the wildtype allele. Hypomethylation of the IAP promoter results in ubiquitously expressed transcript, resulting in yellow coat colour, higher risk of development of tumours and obesity (Jaenisch and Bird, 2003). B. Image of genetically identical week 15 A^{vy} mice. This image demonstrates the phenotypes (colour and obesity) associated with variation in methylation of the A^{vy} allele (Jirtle and Skinner, 2007).

The association between environmentally induced subtle but accumulative epigenetic changes and increased susceptibility to diseases, such as cancer also correlates with the extensive time required for the development of cancers and the late onset of most cancers (Feinberg et al., 2006). For example, in a study conducted on genetically identical monozygotic (MZ) human twins, it was shown that differences in their epigenomic profiles may account for their varying phenotype (i.e. disease) in response to environmental factors over time. Indeed, these epigenetic differences were more apparent between older twins with extremely different lifestyles (i.e. diet, smoking and physical activity). This demonstrates a link between accumulation of epigenetic changes over time and altered phenotype (Fraga et al., 2005). In another study conducted in MZ human twins, it was shown that the promoter region of the *AXIN1* gene was significantly hypermethylated in the twin with a caudal duplication abnormality compared to the unaffected twin and independent controls. This gene has been associated with caudal duplication. Therefore, as no mutation was detected in the *AXIN1* gene of the affected twin, this also indicated that phenotypic differences can arise as a result of epigenetic alterations (Oates et al., 2006). This highlights the importance of epigenetic mechanisms in creating diverse phenotypes within human populations (Whitelaw and Whitelaw, 2006).

In addition, the carcinogenicity of some environmental contaminants such as endocrine disruptors, nickel, cadmium, chromium and arsenic cannot entirely be explained through genetic mechanisms (Martinez-Zamudio and Ha, 2011). Metals can interfere with the activity of DNA methyltransferases, either directly or through production of reactive oxygen species. This leads to an altered DNA methylation profile and subsequent alterations in gene expression (reviewed in Baccarelli and Bollati, 2009). For example, it has been proposed that cadmium (Cd)-induced global DNA hypomethylation may be due to Cd interaction with DNMTs and subsequent interference with their methylation activity. Detoxification of inorganic arsenic (As) is dependent on its enzymatic methylation via the universal methyl donor SAM, thus reducing the amount of available SAM for DNA methylation reactions. This has been linked to global DNA hypomethylation observed in rat liver epithelial cell lines treated with As. In addition, gene-specific hypermethylation (*P53*, *CDKN2A*) has been reported in blood samples collected from individuals exposed to As. Hypermethylation of the *DAPK* promoter was observed in individuals with bladder cancer in areas with known arsenic contamination. Substitution of nickel for magnesium in the DNA triggers chromatin

condensation which is further accompanied by initiation of *de novo* methylation. These changes lead to alteration of DNA methylation patterns at these regions (Arita and Costa, 2009; Baccarelli and Bollati, 2009; Martinez-Zamudio and Ha, 2011). Hence, epigenetically-induced deregulation of key signalling pathways due to exposure to environmental contaminants needs to be taken into consideration when establishing mechanisms of toxicity and tumourigenicity (Feinberg et al., 2006; Martinez-Zamudio and Ha, 2011).

Although it is recognised that the potential for transgenerational epigenetic inheritance is somewhat controversial, the concept that some epigenetic markers (imprinted genes and metastable epialleles) can escape the DNA methylation reprogramming that occurs during embryogenesis means that they can be inherited across generations (Whitelaw and Whitelaw, 2008). Therefore, environmentally-induced epigenetic alterations in the germ line are potentially heritable. This could subsequently lead to inheritance of the environmentally acquired phenotypes. This phenomenon is referred to as transgenerational epigenetic inheritance or epigenetic basis for inheritance of a trait (Anway and Skinner, 2006). For example, several studies have demonstrated that intraperitoneal exposure of gestating outbred Harlan Sprague Dawley female rats to the anti-androgenic endocrine disrupter vinclozolin (VCZ) (100 mg/kg body weight (bw)/day) during critical stages of sex determination (embryonic day (E) 12-E15) induces epigenetically male germ cell transmitted phenotypic characteristics up to at least the F₃ generation in male rat offspring, such as increased spermatogenic cell apoptosis, decreased sperm motility and numbers, prostate abnormalities, tumours and renal lesions. The reproducibility and frequency of the VCZ-induced phenotypes (i.e. rate of tumours) and identification of genes with altered DNA methylation in the effected individuals compared to controls, indicated that mutations are not the most likely cause of this abnormality (Anway et al., 2005; Anway et al., 2008a; Anway et al., 2008b; Anway and Skinner, 2008; Anway et al., 2006; Stouder and Paoloni-Giacobino, 2010). However, It has been demonstrated that the effect and epigenetic inheritance of VCZ is highly dependent on dose (Skinner et al., 2010), animal strain (Inawake et al., 2009) and route of exposure (Schneider et al., 2008). Oral administration of VCZ (100 mg/kg bw/day) in outbred Wistar rats (Schneider et al., 2008) and intraperitoneal injection of VCZ (100 mg/kg bw/day) in inbred CD-Sprague Dawley rats (Inawake et al., 2009) during the sex determination stage failed to induce inherited phenotypic effects.

The environmentally-induced epigenetic modifications have the potential to cause change in the expression of genes rather than change in DNA sequences of genes. As a result key genes with altered expression as a consequence of response to environmentally-induced epigenetic changes can act as epigenetic biomarkers or “fingerprints”, for early detection of disease and environmental exposures (this concept is further discussed in Chapter 6) (Jirtle and Skinner, 2007). However, it has to be noted that this does not mean that genetically-induced modifications, caused by environmental factors (i.e. chemical contamination), are not important and can be discarded. But rather, both epigenetic and genetic mechanisms and the possibility of epigenetic and genetic polymorphisms need to be taken into consideration (Bollati and Baccarelli, 2010). Therefore, in the same way that genetic polymorphisms can influence the susceptibility of individual organisms to toxicity and disease, differences in the epigenome that emerge throughout the life-time of an individual may also have the same effects (Christensen et al., 2009) (Figure 1.14).

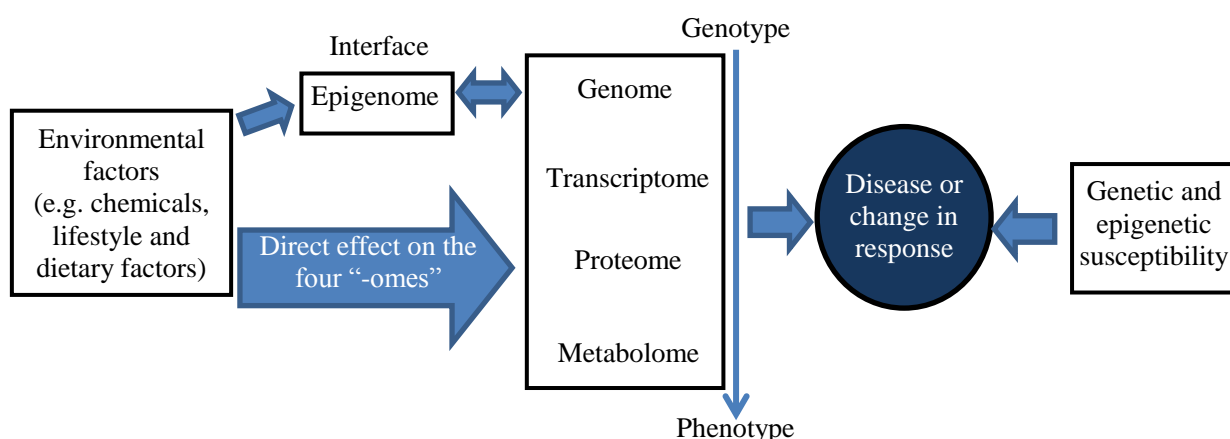


Figure 1.14. The interactions between environmental factors and the various “-omes”, including the epigenome. There is an interactive relationship between the five –omes. Environmental factors can directly affect the traditional –omes (genome, transcriptome, proteome and metabolome) or induce changes to the epigenome. The epigenome acts as an interface between the genome and the environment. Alterations to the epigenome have the potential to induce changes in gene expression levels and subsequently alter the phenotype. In addition to environmentally-induced changes in responses (e.g. higher risk of disease), both genetic and epigenetic susceptibility of an individual should be considered. However, both higher genetic and epigenetic susceptibilities of an individual to certain phenotypes (e.g. diseases) could have environmentally-induced origins.

1.3.5.3. Two models of tumourigenesis: The classic multistep model initiated by mutation and the epigenetic progenitor model

Originally cancer was viewed as a genetic disease initiated by mutations. However, it is now known that cancers are a consequence of accumulation of both genetic and epigenetic alterations. These events complement each other to promote tumourigenesis by activation and inactivation of the genes that are associated with neoplastic transformation (Pogribny, 2010; Herceg and Vaissiere, 2011). Furthermore, the need for identifying a common event between different types of tumours and establishing the order of events in neoplastic transformation, has led to the proposal of various hypotheses and models of cancer development. Below two models are briefly described: 1. The multistep model of tumourigenesis, which until recently has been the universally accepted model of tumourigenesis 2. The epigenetic progenitor model of tumourigenesis. The latter has been proposed in the light of a better understanding of the significance of epigenetic mechanisms and the role of epigenetics as an interface between environmental factors and genome.

1. Multistep model of tumourigenesis (initiation, promotion and progression)

This well-defined classic model is comprised of an initiation step by a genotoxic compound, a promotion and a progression step (Weinberg, 1989). During the first step, exposure to a genotoxic compound -chemicals that induce cancer by directly damaging the DNA (Mally and Chipman, 2002)- leads to formation of an irreversible mutation (Barrett, 1993; Klaunig et al., 2000; Moggs et al., 2004). Thus, during DNA replication the modified DNA sequence is passed on to the next generation of the cells and becomes permanently fixed in the genome (Klaunig et al., 2000). Increased proliferation and inhibition of apoptotic signals results in clonal expansion of the initiated cell. Growth advantages of a group of daughter cells due to further epigenetic and genetic changes result in their selection over other initiated cells. This step is referred to as promotion (Barrett, 1993; Klaunig et al., 2000). Finally, during the progression stage, further changes of a sub-section of the promoted cells result in their ability to overcome their dependency on the other cells and upon extra-cellular signals. This results in metastasis of tumour cells and invasion of other tissues (Figure 1.15) (Moggs et al., 2004; Klaunig et al., 2000).

The model described above relies on a genotoxic compound or event to initiate tumourigenesis by irreversibly damaging and mutating the DNA (Moggs et al., 2004), and relies on the participation of non-genotoxic mechanisms only after the initial step. This group contributes to tumourigenesis by promoting cell proliferation and inhibiting apoptosis. This imbalance between proliferation and apoptosis caused by non-genotoxic compounds, can be achieved through different mechanisms, such as alterations of growth signals, gene expression and gap junctional intercellular communication (Klaunig et al., 2000).

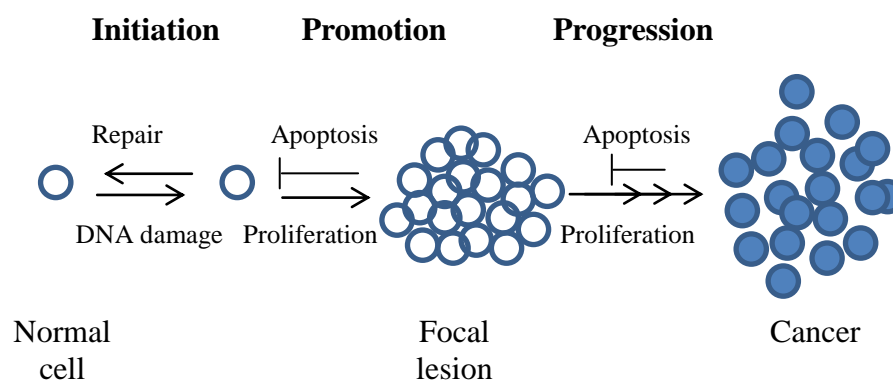


Figure 1.15. Multistage process of carcinogenesis. Sufficient un-repaired damage to the DNA leads to formation of an initiated cell. Further promotion of the initiated cell by mechanisms such as increase in proliferation and inhibition of apoptosis results in selective clonal expansion of the initiated cell and formation of a cluster of modified cells, known as a “focal lesion”. Finally, cell growth and further alterations of the focal lesion lead to formation of a fully formed cancer (Klaunig et al., 2000).

2. Epigenetic progenitor model of tumourigenesis

The basis of tumourigenesis is the irreversible loss of control of a cell in terms of regulation of its genetic material, transcription and proliferation levels (Moggs et al., 2004). This results in generation of an undifferentiated, stem cell-like, immortal cell with unlimited proliferation abilities (Moggs et al., 2004; Sharma, et al., 2010; Herceg and Vaissiere, 2011; Pogribny, 2010; Feinberg et al., 2006). Better understanding of the function of epigenetic mechanisms within a cell and the characteristics of tumour cells led to the knowledge that a stable epigenetic change, such as deregulation of the central epigenetic control machinery, could

initiate a neoplastic transformation (Moggs et al., 2004; Sharma et al., 2010). Knowing that epigenetic mechanisms, such as DNA methylation, can greatly influence gene expression and differentiation of the cells has given rise to the epigenetic progenitor/stem cell model of tumourigenesis (Feinberg et al., 2006; Moggs et al., 2004). It is important to mention that this model does not disagree with the importance of genetic alterations such as mutation in gatekeeper genes (gatekeeper genes are genes that can directly regulate tumour growth and therefore, are important at early stages of tumour development) and chromosomal rearrangement in tumourigenesis. Indeed, mutations in the gatekeeper genes are necessary in the early stages of tumour development. For example, mutation in the gatekeeper *APC* gene and translocation and formation of the oncogenic BCR-ABL (break cluster region-V-abl Abelson murine leukaemia viral oncogene homolog 1) gene fusion have been recognised as key steps in development of colorectal cancer and chronic myelocytic leukaemia, respectively. However, chromosomal rearrangement and changes in the expression of gatekeeper genes can also be induced by epigenetic mechanisms and have the same effects as mutations (Feinberg, 2004; Feinberg et al., 2006).

The epigenetic progenitor model proposes that cancer is initiated through epigenetic changes in progenitor cells much earlier than originally thought. Three steps have been proposed for this model. The first step is the main difference between the two models. 1. Epigenetic disruption of the progenitor cells 2. Mutations in gatekeeper genes 3. Genetic and epigenetic plasticity (Feinberg et al., 2006). These stages are briefly described below and are shown in Figure 1.16.

1. Epigenetic disruption of progenitor cells. This step involves epigenetic alterations of “tumour-progenitor genes” within progenitor cells of a given tissue (somatic stem cells) and formation of a cancer-primed population of progenitor cells. Tumour-progenitor genes are defined as a group of genes that are responsible for characteristics of stem cells (i.e. genes involved in differentiation, intercellular communication, proliferation and genes with wide genetic and epigenetic consequences), for example, *IGF2*, connexins and tumour suppressor gene *APC* which are all known to be epigenetically altered in tumours (Feinberg et al., 2006; Moggs et al., 2004). Therefore, the basis of this step relies on two concepts, the stem cell origin of cancers and the occurrence of epigenetic changes, with the ability to alter biological functions of the cells, prior to genetic alterations.

The similar characteristics of somatic stem cells and cancer cells have led to the proposal of stem cell origin of cancers (Pardal et al., 2003; Feinberg et al., 2006). Without doubt both cell types can self-renew and differentiate, though both actions occur in a more controlled and effective manner in somatic stem cells compared with cancer cells. For example, although poorly and abnormally differentiated, myeloid leukaemia cells can differentiate to form diverse blood cell lineages. Also, medulloblastomas, a type of brain tumour, can differentiate to form neuron and glia-like cells. Furthermore, in many tumour types, such as breast cancer, glioblastoma and acute myeloid leukaemia, two distinct groups of cells, stem-like and more differentiated tumour cells are detectable based on cell surface markers (Pardal et al., 2003), therefore giving rise to the stem cell/progenitor origin of cancers.

As discussed in section 1.3.4.3, epigenetic mechanisms are fundamental in differentiation of stem cells. Hence, their disruption can alter the transcription of genes and cause chromosomal instability and provide these cells with growth and survival advantages (stem cell properties). These changes alter the balance between differentiated, committed cells and progenitor cells. In addition, detection of epigenetic abnormalities in normal tissues in the absence of mutations and prior to development of tumours, strengthens the concept of involvement of epigenetic mechanisms and abnormalities in initiation of cancers (Feinberg et al., 2006; Hecceg and Vaissiere, 2011). For example, DNA hypomethylation has been detected in premalignant stages of gastric carcinoma (Cravo et al., 1996) and it induces aggressive T-cell lymphoma in mice (Gaudet et al., 2003). Furthermore, in colorectal cancer DNA hypomethylation can precede mutational changes, as it has been detected in normal-appearing tissues of patients harboring adenomas prior to neoplastic transformation (Choi and Mason, 2002). LOI of the *IGF2* growth factor has been associated with 5-fold increase in risk of colorectal cancer. This modification has also been found throughout the non-cancerous colonic epithelium of patients with LOI-associated colorectal cancer (Cui et al., 2003).

Several other pieces of evidence such as the reversibility of tumour growth at very early stages of tumourigenesis and detection of a uniform DNA hypomethylation in all the cells of a tumour mass tends to infer that epigenetic changes occur prior to mutations (Feinberg et al., 2006).

2. Initiating mutations in gatekeeper genes. This step is similar to the classic model with the exception that genetic changes in gatekeeper genes are not seen as the initiating step and occur in primed cells.

3. Genetic and epigenetic plasticity. Further epigenetic and genetic changes lead to formation of a fully formed tumour with the phenotypic characteristics of tumour cells (Feinberg, 2004; Feinberg et al., 2006). Epigenetic-induced instability and the role of epigenetics in cancer are described in section 1.3.5.1. The epigenetic progenitor model of tumourigenesis is shown in Figure 1.16.

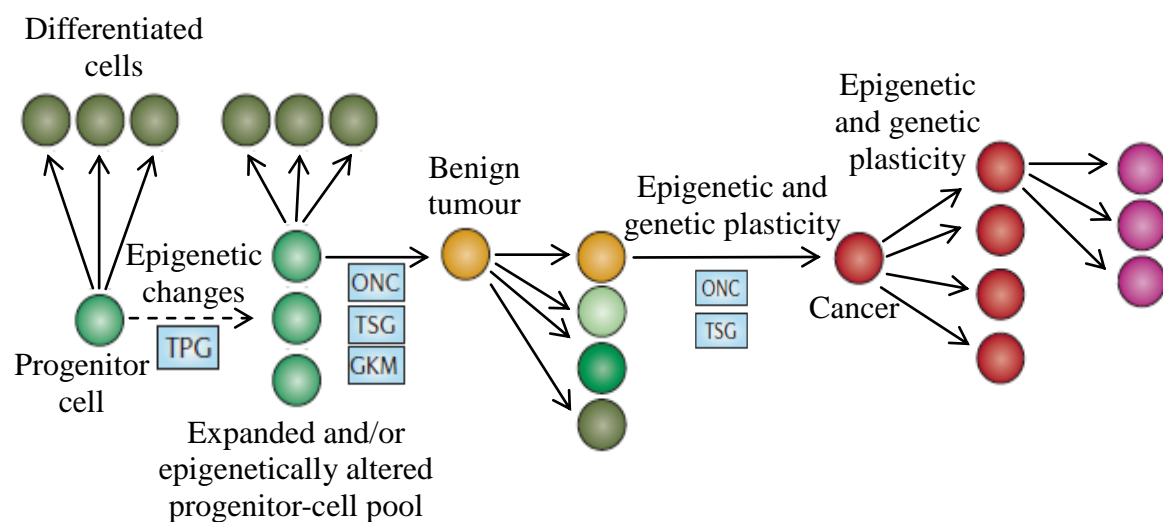


Figure 1.16. Epigenetic progenitor model of tumourigenesis. Cancer is initiated by epigenetic disruption of progenitor cells. This results in generation of a tumour-primed, abnormal population of progenitor cells that can maintain their stem-cell like properties. This is achieved by deregulation of tumour-progenitor genes (TPG). The second step is mutation in gatekeeper genes (GKM) which includes oncogenes (ONC) and tumour suppressor genes (TSG). Finally, further epigenetic and genetic instabilities lead to formation of fully formed tumour cells with features associated with cancer (Feinberg et al., 2006).

A better understanding of the epigenetic mechanisms and their role in development of disorders and acceptance of the epigenetic model of tumourigenesis has several outcomes. This affects our understanding of the biology of cancer, the identification of new targets for design of antineoplastic drugs (e.g. DNA methyltransferase inhibitors and HDAC inhibitors),

the role of environment in development of disorders (i.e. epigenetic memory) and chemical safety assessment (Moggs et al., 2004; Feinberg et al., 2006). This model implies that epigenetic changes during tumourigenesis occur at an early stage, prior to the stage that is currently recognised as benign pre-neoplastic lesions. Therefore, recognising the early reversible epigenetic changes as early biomarkers of prevention or development of cancer are extremely valuable (Feinberg et al., 2006).

1.4. DNA methylation and its implication in marine biology

Although the role of epigenetic mechanisms in regulation of transcription, response to internal and external environmental factors and human health are becoming more apparent, this area is under-studied in aquatic biology, especially in relation to disease. If chronic exposure to environmental contaminants can cause epigenetic changes associated with disease in model organisms; there is a high probability that the same mechanisms are implicated in the development of disease in aquatic species and are affecting their responses to environmental factors. This becomes more important when considering that in current biomonitoring protocols and in procedures used for evaluation of health of aquatic species and environmental quality, long term exposures to chemical contaminants (especially in the case of persistent organic pollutants) and their effects upon epigenetic mechanisms are not considered at all. Taking into account that exposures to environmental contaminants can induce epigenomic changes in key signalling pathways, it becomes extremely important to consider the role of epigenetic mechanisms in the diseases of aquatic species. Furthermore, as chemicals can alter the epigenome with subsequent changes in the transcription, it may be possible to identify a subset of genes with altered DNA methylation (biomarkers) unique to each category of chemicals. Identification of these biomarkers can potentially be used in bimonitoring, early detection of exposure to chemical contaminants and chemical risk assessments (Moggs et al., 2004; Prins, 2008). However, it has to be noted that some epigenetic modifications are reversible and some occur normally. Hence, when identifying biomarkers it is important to distinguish between adverse and non-adverse epigenetic signals and a change in DNA methylation does not necessarily represent a change in phenotype or toxicity (Moggs et al., 2004).

1.4.1. Cancer in fish

1.4.1.1. Use of fish in carcinogenicity studies

Fish have been used in carcinogenicity studies, as biomonitoring species to provide information on the quality of the environment and the health of the population of fish species and as alternative non-mammalian cancer models. Indeed, their high sensitivity to a variety of classes of carcinogens, tumour promoters, low cost of maintenance and short life cycles makes them ideal test species for both environmental monitoring and investigation of molecular mechanisms of tumourigenesis (Small et al, 2010; Bailey et al., 1996). In addition, although the normal structures of some organs are different between mammals and fish, tumours in fish are histopathologically similar to the equivalent tumours in mammals. This allows the same tumour type classification in fish as in mammals and their widespread use in research (Masahito et al., 1988; Grabher and Look, 2006). For instance, rainbow trout (*Oncorhynchus mykiss*) has been used for biomonitoring of contamination with environmental chemical carcinogens for the last four decades. Their low maintenance cost, ultra-sensitivity at very early stages of life and sensitivity to different classes of carcinogens, in addition to their well characterised tumour pathology, has led to their extensive use in research. As an example, investigations into the cause of high rates of liver tumour incidence in rainbow trout collected from the Pacific Northwest led to the discovery of aflatoxin B1 (AFB1) as a human hepatocarcinogen. Further investigations into AFB1-induced hepatocellular and cholangiocellular carcinomas in trout, revealed that the *c-k-ras* gene was mutated in these tumours, similar to rat liver tumours (Bailey et al., 1996). Another example of species of fish used in carcinogenicity studies is medaka (*Oryzias latipes*). Medaka is commonly used in carcinogenicity studies. The effects of more than 30 common carcinogens and their links with formation of hepatocellular carcinoma are well characterised in this fish species. A number of these carcinogens include AFB1, *N*-diethylnitrosamine and methylazoxymethanol acetate (Bailey et al., 1996). It was also demonstrated that the functional regions of the *retinoblastoma* gene is conserved between humans and medaka and mutations in this gene can lead to formation of retinoblastoma. These studies as well as highlighting the importance and usefulness of fish species in testing for carcinogenicity, demonstrate the potential use of fish as models for studying cancers in humans (Rotchell et al., 2001b).

1.4.1.2. Common dab (*Limanda limanda*) and environmental carcinogenicity studies

Environmentally-induced tumours in fish collected from their natural habitat are mainly observed in bottom-dwelling fish. This has been linked with their high levels of exposure to sediment-associated chemical carcinogens (Masahito et al., 1988). For example an unusually high prevalence of liver tumours, with some sampling sites exceeding 20%, has been reported in the flatfish dab captured from the waters around the UK compared with the frequencies considered to represent a background prevalence of the disease (NMMP, second report, 2004; Lyons et al., 2006). Other bottom-dwelling flatfish with reported high levels of hepatocellular carcinoma and cholangiocarcinoma tumours include English sole (*Pleuronectes vetulus*) and winter flounder (*Pleuronectes americanus*) sampled from Boston Harbor, USA. Although tumours are mainly observed in bottom-dwelling flatfish, they are not restricted to these species. For example, tumours have been reported in a wide variety of species including brown bullhead (*Ictalurus nebulosus*) in the rivers entering Lake Erie and also in white perch (*Morone americana*) from Chesapeake Bay (Masahito et al., 1988).

As a result, monitoring the disease status of dab and Atlantic cod (*Gadus morhua*) in offshore waters and European flounder (*Platichthys flesus*) in inshore regions and estuaries is now established as part of the biomonitoring procedures set up by the International Council for the Exploration of the Seas (ICES). These species are monitored for both internal (i.e. foci of cellular alterations (FCA) and malignant tumours) and external signs of disease (NMMP, second report, 2004). Therefore, the presence of tumours is used as a possible indicator of chemical contaminants at levels which cause adverse health effects and serve as indicators of the health and quality of the marine environment and fish population (NMMP, second report, 2004; Masahito et al., 1988). In addition to chemical induction, it is possible that biological agents such as viruses are causative of tumours. This was shown in the case of viral-induced neurofibromatosis in bicour damselfish (*Stegastes partitus*) collected from waters around Florida, USA (Schmale et al., 2002).

Common dab (Figure 1.17) is similar to the other flatfish species living in close proximity to the ocean floor. Due to its diet of sediment-dwelling invertebrates it can be exposed to relatively high levels of sediment-associated chemicals. It is therefore a useful species for monitoring of the bioaccumulation of organic compounds and environmental carcinogenicity research. Liver pathology in dab, including cancer and pre-neoplastic lesions, is used as an

indicator of the biological effects of contaminants on the marine environment. Recently, attention has focussed on dab caught from UK waters as part of the UK Clean Seas Environmental Monitoring Programme (CSEMP). This is due to the detection of an unusually high prevalence of liver tumours in dab, with some UK sites exceeding 20% (NMMP, second report, 2004; Lyons et al., 2006). Lesions seen in livers of flatfish dab can be separated into 5 main categories; non-specific inflammatory responses, non-neoplastic lesions, FCA, benign tumours and malignant tumours. (NMMP, second report, 2004; Stentiford et al., 2009; Small et al., 2010; Southam et al., 2008; Lyons et al., 2006). Liver is the main organ used for investigation and detection of morphological alterations and detection of tumours in fish collected for biomonitoring. Liver and kidney are the two main organs for xenobiotic metabolism. Based on studies conducted on the livers of rainbow trout, flatfish and zebrafish it has been shown that, similar to humans, fish liver expresses the main phase I and biotransformation enzymes, e.g. cytochrome P450s (CYP), required for activation of the pro-carcinogens and metabolism of compounds and therefore is the main target site for chemically-induced adverse effects. Furthermore, first pass metabolism, in addition to high levels of CYPs, increases the suitability of the liver as a biomonitoring organ. For example, tumours are developed in the liver of rainbow trout in response to exposure to CYP1A-inducing compounds such as 3-methylcholanthrene. Induction of CYP1A is used as a biomarker of exposure to polycyclic aromatic hydrocarbons (PAHs) (Bailey et al., 1996; Bragigand et al., 2006; Sheader et al., 2006). However, different types of tissue can be used for monitoring of different toxicological endpoints. For example skin and gills in fish are extremely useful for monitoring the visible morphological changes as they are directly exposed to the environment. In contrast, biofluids offer a less invasive method for biomonitoring (Lindesjoo and Thulin, 1994; Ward et al., 2006).

However, what is most interesting is that the levels of tumours in dab collected from offshore regions are higher than the levels of tumours detected in its close relative, the European flounder, which is collected from inshore areas. Theoretically, European flounder should be exposed to higher levels of contaminants due to their closer proximity to coastal regions and sources of anthropogenic pollution. So far the causative agents and the molecular mechanisms of these tumours remain unclear, especially in regards to the involvement of environmentally-induced epigenetic changes and their role in dab tumourigenesis. Better understanding of the pathways altered in these tumours and better characterisation of these tumours at molecular

levels will help to clarify the driving factors behind these tumours. In addition, this may help to identify biomarkers for identification of hazardous environments and early indicators of health problems in fish. As a result, two recent studies by Small et al., (2010) and Southam et al., (2008) attempted to analyse dab liver tumours at the transcriptomic and metabolomic levels, respectively. Both studies provided preliminary data in terms of tumour characterisation and the identification of pathways that are altered in these tumours (i.e. energy pathways, protein synthesis). In the first study, a cDNA microarray designed for the close relative of dab, the European flounder, was used. Although this microarray is limited in the number of the genes presented on this platform (27,648 features –derived from 13,824 clones spotted in duplicate. The clones were derived from a redundant flounder cDNA library and represent approximately 3336 unique sequences), nevertheless the authors showed that genes involved in protein synthesis and vitellogenin, the egg precursor protein, were significantly up-regulated in dab tumours in comparison with surrounding non-tumour liver tissues. As vitellogenin up-regulation in fish is used as a biomarker of exposure to endocrine disrupting chemicals, the authors concluded that these chemicals may be involved in the development of these tumours. However they failed to identify a possible mechanism/evidence for induction of liver tumours by these chemicals in dab.

In the second paper (Southam et al., 2008) the authors used one-dimensional ^1H NMR spectroscopy approach for identification of the metabolites whose concentrations were significantly altered in dab tumours. One of their findings was that concentrations of two metabolites corresponding to choline and betaine were altered in dab tumours. As these metabolites are involved in the one-carbon cycle; therefore, they concluded that possibly the DNA methylation pathway is disrupted in these tumours. However, this work required further investigation of the key metabolites in the one-carbon pathway (SAM and SAH) and required evidence of altered DNA methylation at gene levels.

Furthermore, both studies mainly focused on characterisation of the tumours and were limited to tumours and apparently healthy tissues surrounding these tumours. Including completely healthy liver tissues from these fish species as a third category may help to identify the involvement of environmental factors and the possibility of higher susceptibility of the tumour bearing fish to chemicals. However, considering the challenges of working with non-model, un-sequenced species; both studies have provided significant insight into the pathways that are altered in these tumours.



Figure 1.17. A picture of the flatfish dab (*Limanda limanda*).

1.4.1.3. Zebrafish (*Danio rerio*) as a laboratory cancer model

The initial use of zebrafish in developmental biology dates back to a few decades ago. The transparency and *ex-utero* development of zebrafish embryos made this species an ideal model for studies on developmental biology. Its use has now expanded to different aspects of biosciences, including cancer biology and it is now being recognised as a convenient model for studying human tumourigenesis alongside more traditionally used models, such as mice, rats and nonhuman primates (Feitsma and Cuppen, 2008; Lawrence et al., 2009). Several factors make zebrafish a suitable model for cancer research. These include the sequenced genome, sensitivity to a variety of carcinogens, low cost of maintenance, short reproductive cycle, possibility of field studies and portability, potential for lifetime bioassays, transplantation of fluorescent mammalian cells and ease of forward and reverse screening (Bailey et al., 1996; Spitsbergen et al., 2000; Feitsma and Cuppen, 2008). In addition, fundamental concepts in development of tumours in humans, such as genome instability, tumour invasion and progression, presence of cancer stem cells, tumour suppressor genes and

oncogenes have also been recognised in the processes that lead to development of tumours in zebrafish (Feitsma and Cuppen, 2008).

In addition to the significant histopathological similarities between zebrafish tumours and corresponding human tumours (Lam and Gong, 2006), recently it has been shown that transcription profiles of the tumours between the two phylogenetically distant species are also similar (Lam et al., 2006). Indeed, gene expression signatures in chemically-induced liver tumours in zebrafish show clear similarities to those in human tumours. Thus, genes involved in cell cycle/proliferation, DNA replication and repair, apoptosis and genes with liver-specific function were found to be deregulated in both human and zebrafish liver tumours. Furthermore, pathways such as Wnt- β -catenin and Ras-MAPK which are commonly distorted in human liver cancers, especially in hepatocellular carcinoma (HCC), are also altered in zebrafish liver tumours (Cha and DeMatteo, 2005; Lam et al., 2006).

Finally, although there are many benefits in the use of zebrafish for cancer research, there are still many unknown factors associated with the use of zebrafish (Feitsma and Cuppen, 2008) or any other fish species both as human cancer models and for environmental carcinogenicity studies. The role of epigenetics in fish tumourigenesis and environmentally-induced changes is one of these factors. So far no research has been conducted to investigate the involvement of epigenetic mechanisms in the development of tumours in fish. Indeed, consideration of epigenetic factors as one of the key components involved in human tumourigenesis cannot be disregarded in any model that it is currently in use for studying neoplasia in humans. This further highlights the need for studying epigenetic mechanisms in these species.

1.5. Aims

The molecular mechanisms and the reasons for high prevalence of environmentally-induced tumours in the flatfish dab are unknown, especially in relation to the balance between epigenetic and genetic factors. In addition, previous studies in our laboratory indicated that commonly mutated genes in human cancers (e.g. *k-ras* and *h-ras*) were not mutated in liver tumours of the close relative of dab, the European flounder (Rotchell et al., 2001a; Franklin et al., 2000). Therefore, based on the preliminary findings of changes observed in the one-carbon cycle in tumours of these fish (Southam et al., 2008), the importance of epigenetic mechanisms in development of cancers in humans, the role of epigenetics as an interface

between environment and genome and the limited information on the significance of epigenetic mechanisms in fish tumourigenesis, it was highly appropriate to investigate the possibility of involvement of epigenetic mechanisms in induction of and in the features of these tumours. However, as dab is not sequenced, conducting DNA methylation studies are extremely challenging. As a result, to identify a role for involvement of epigenetic mechanisms in fish tumourigenesis we tested the hypothesis that DNA methylation was altered in chemically induced fish liver tumours using zebrafish, which lent themselves to effective analyses based on genome sequence availability.

The aims of this project were:

1. To test the involvement of DNA methylation mechanisms in chemically-induced hepatocellular carcinoma (HCC) tumours in zebrafish (*Danio rerio*).
2. To overcome the challenges encountered when studying DNA methylation in non-model, un-sequenced organisms.
3. To characterise hepatocellular adenoma (HCA) tumours of dab (*Limanda limanda*). This was investigated at two levels, DNA methylation and gene transcription. These studies were conducted with the aim of better understanding of the molecular mechanisms and molecular features of these tumours and developing an evidence-based mode-of-action from the original hypothesis that epigenetic factors are involved in the high prevalence of naturally occurring liver tumours in flatfish dab.
4. To conduct the first targeted metabolomic study of the one-carbon cycle in fish tumours with the aim of testing the premise that this pathway is altered leading to modulation of the epigenome in dab HCA tumours.

The overall aim of this project was to determine if DNA methylation profile was significantly altered in fish liver tumours compared to healthy fish liver.

Chapter 2

Materials and Methods

2.1. Chemicals

All chemicals were obtained from Sigma-Aldrich Chemical Company (UK) unless otherwise stated.

2.2. Test organisms

2.2.1. Zebrafish (*Danio rerio*)

2.2.1.1. Chemical induction of hepatocellular carcinoma in zebrafish liver

Exposure of zebrafish to carcinogen and extraction of DNA from zebrafish livers were carried out by our collaborators at the National University of Singapore. The experimental procedure was approved by their Institutional Animal Care and Use Committee (Lam et al., 2006).

Briefly, three-week post-fertilisation zebrafish fry (100_150) were treated with 7, 12-dimethylbenz[α]anthracene (DMBA, 0.75 parts per million (ppm)) in dimethyl sulfoxide (DMSO) or DMSO (vehicle control) for 24 hours and treatment was repeated once, 2 weeks later, for another 24 hours with 1.25 ppm DMBA or DMSO (The chemical concentrations are ppm in fish tank water). Treated fish were rinsed three times in fresh water and transferred into new tanks for maintenance. Fish were sampled 6_10 months after DMBA exposures. The tumour samples used for the present study were all larger than 3 mm in diameter. Liver tumours were sampled for histopathological diagnosis (Figure 2.1). In the zebrafish study (Chapter 3), healthy male livers from the vehicle control exposure (n=4) and HCC from the DMBA exposed group (n=4, male) were used.

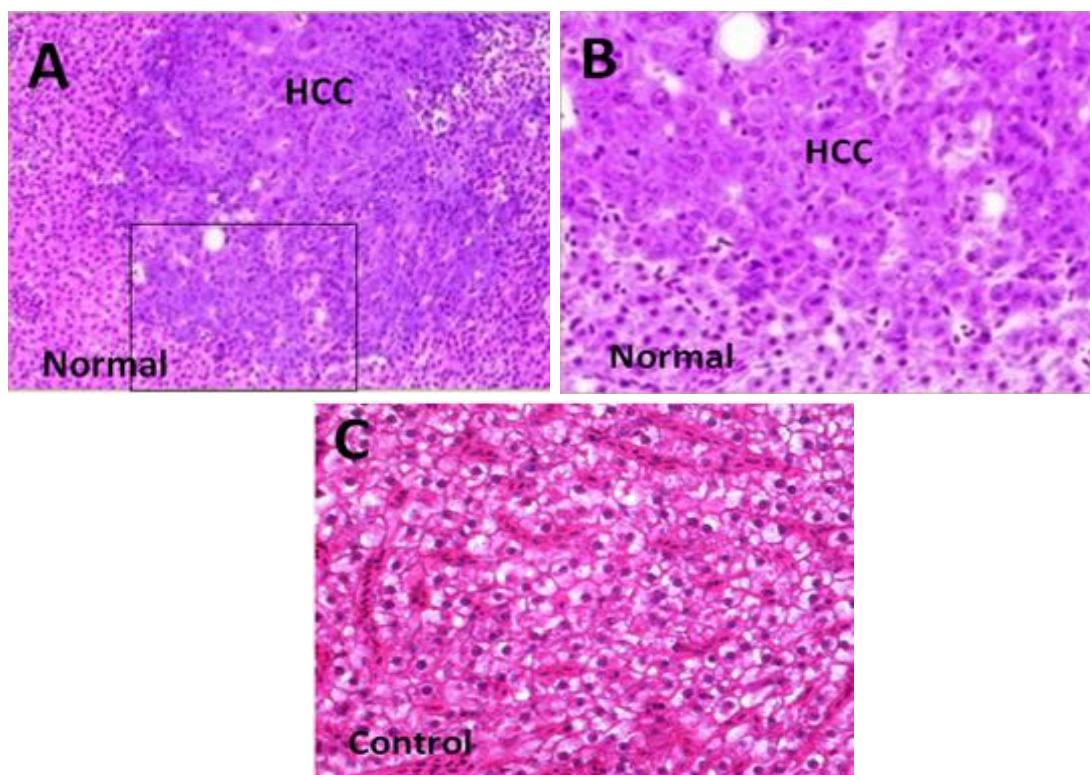


Figure 2.1. Histopathology images of zebrafish HCC and healthy zebrafish liver.

A. Zebrafish liver section with hepatocellular carcinoma (HCC) invasion into surrounding normal liver tissue. HCC and surrounding normal tissues are labelled. B. Enlarged image of the section indicated with a box in image A. This image demonstrates the border between HCC and normal tissue. C. Normal control liver section. Normal zebrafish hepatocytes are typically organised in two-cell thick plates and are regular throughout the whole liver. Carcinomas lose this plate architecture completely and are reorganised in typical patterns, such as a trabecular pattern with several-cell thick irregular trabeculae, a glandular pattern with a central clear space surrounded by one-cell layer neoplastic hepatocytes, and a large sheet of neoplastic cells without any recognisable pattern. Carcinoma cells are cuboidal with centrally localised nuclei of variable sizes.

2.2.1.2. DNA extraction from zebrafish liver samples

Zebrafish liver samples were placed in RNAlater (Ambion, Austin, Texas). DNA samples were extracted from the tissue samples using Trizol reagent (Invitrogen, UK) according to the manufacturer's instructions.

Zebrafish tissues were homogenised in Trizol (1 ml of Trizol/100 mg of tissue) and incubated at 30°C for 5 minutes. Chloroform (200 µl/1ml of Trizol) was added to the samples, mixed by vortexing and incubated at 30°C for 15 minutes. Following incubation samples were

centrifuged at 13400 x g for 15 minutes at 4°C. This resulted in separation of the samples to an aqueous layer containing RNA, an interphase layer and an organic layer containing DNA and proteins. The aqueous layer was removed for purification of RNA samples. The RNA samples were not used in this thesis. Absolute ethanol (300 µl/1ml of Trizol) was added to the organic layer containing the DNA sample. Samples were gently mixed, incubated at 30°C for 3 minutes followed by centrifugation at 13400 x g for 5 minutes at room temperature. Supernatant was removed and the DNA pellet was washed twice with 0.1M sodium citrate (1ml sodium citrate/1ml of Trizol). Samples were incubated at 30°C in wash solution for 20 minutes with occasional mixing. DNA samples were centrifuged at 13400 x g for 5 minutes at room temperature. Supernatant was removed and DNA samples were air dried and stored at -20°C for subsequent use in the experiments. The zebrafish exposures, histopathology confirmation of types of tumours and extraction of DNA was carried out by our collaborators at the National University of Singapore. The complete detail of their experimental approach is described in their publication (Lam et al., 2006).

2.2.2. Dab (*Limanda limanda*)

2.2.2.1. Collection of dab livers

As part of the CSEMP, several hundred dab flatfish were captured from sampling sites in the Irish Sea, the North Sea and the Bristol Channel (Table 2.1) during June and July 2007, 2008 and 2009 using 30 minute tows of a standard Granton trawl by the Centre for Environment, Fisheries and Aquaculture Science (Cefas, Weymouth, UK). Dab were immediately removed from the catch and placed into flow-through tanks containing aerated seawater. Fish were assessed for external diseases, sacrificed and livers were visually assessed for the presence of macroscopic lesions (nodules). Methods described by Feist et al., (2004) were used for dissecting the nodules and the healthy liver tissues. Sections of the collected samples were immediately snap frozen in liquid nitrogen and remained frozen at -80°C for use in DNA methylation, transcriptomic and metabolomics studies and the remainder of the samples were fixed in 10% neutral buffered formalin for 48 hours for histopathology. Histopathology was conducted by Cefas.

Fish ID	Type of tissue	Location	Sea	Sex	Year collected	Type of study
RA09065-830	HCA	South Cardigan Bay	Irish Sea	Female	2009	1, 2, 3, 4
RA09065-830	ST	South Cardigan Bay	Irish Sea	Female	2009	1, 2, 3, 4
RA09065-959	HCA	Inner Cardigan Bay	Irish Sea	Female	2009	1, 2, 4
RA09065-959	ST	Inner Cardigan Bay	Irish Sea	Female	2009	1, 2, 4
RA08044-302	HCA	North Cardigan Bay	Irish Sea	Female	2008	1, 2, 4, 5
RA08044-302	ST	North Cardigan Bay	Irish Sea	Female	2008	1, 2, 4, 5
RA08044-289	HCA	St Bee's Head	Irish Sea	Female	2008	1, 2, 4, 5
RA08044-289	ST	St Bee's Head	Irish Sea	Female	2008	1, 2, 4, 5
RA08045-30	HCA	St Bee's Head	Irish Sea	Female	2008	1, 2, 4, 5
RA08045-30	ST	St Bee's Head	Irish Sea	Female	2008	1, 2, 4, 5
RA09065-1283	HCA	St Bee's Head	Irish Sea	Female	2009	1, 2, 4
RA09065-1283	ST	St Bee's Head	Irish Sea	Female	2009	1, 2, 4
RA08044-96	ST	Red Wharf Bay	Irish Sea	Female	2008	1, 2, 4, 5
RA08044-96	HCA	Red Wharf Bay	Irish Sea	Female	2008	1, 2, 4, 5
RA08045-10	HCA	Red Wharf Bay	Irish Sea	Female	2008	1, 2, 4, 5
RA08045-10	ST	Red Wharf Bay	Irish Sea	Female	2008	1, 2, 4, 5

Continued from previous page

RA09065-886	HCA	South Cardigan Bay	Irish Sea	Female	2009	1, 2, 4
RA09065-886	ST	South Cardigan Bay	Irish Sea	Female	2009	1, 2, 4
RA09065-904	HCA	South Cardigan Bay	Irish Sea	Female	2009	1, 2, 4
RA09065-904	ST	South Cardigan Bay	Irish Sea	Female	2009	1, 2, 4
RA09044-381	HCA	Lundy	Bristol Channel	Female	2009	5
RA09044-381	ST	Lundy	Bristol Channel	Female	2009	5
RA08045-43	HCA	Indefatigable Bank	North Sea	Female	2008	5
RA08045-43	ST	Indefatigable Bank	North Sea	Female	2008	5
RA08045-49	HCA	Indefatigable Bank	North Sea	Female	2008	5
RA08045-49	ST	Indefatigable Bank	North Sea	Female	2008	5
RA07045-9	H	South Cardigan Bay	Irish Sea	Female	2007	1, 2, 4
RA07045-533	H	North Cardigan Bay	Irish Sea	Female	2007	1, 2, 4
RA07045-602	H	Lundy	Bristol Channel	Female	2007	1, 2, 4
RA07045-607	H	Lundy	Bristol Channel	Female	2007	1, 2, 4
RA07045-106	H	Red Wharf Bay	Irish Sea	Female	2007	1, 2, 4
RA07045-107	H	Red Wharf Bay	Irish Sea	Female	2007	1, 2, 4

Continued from previous page

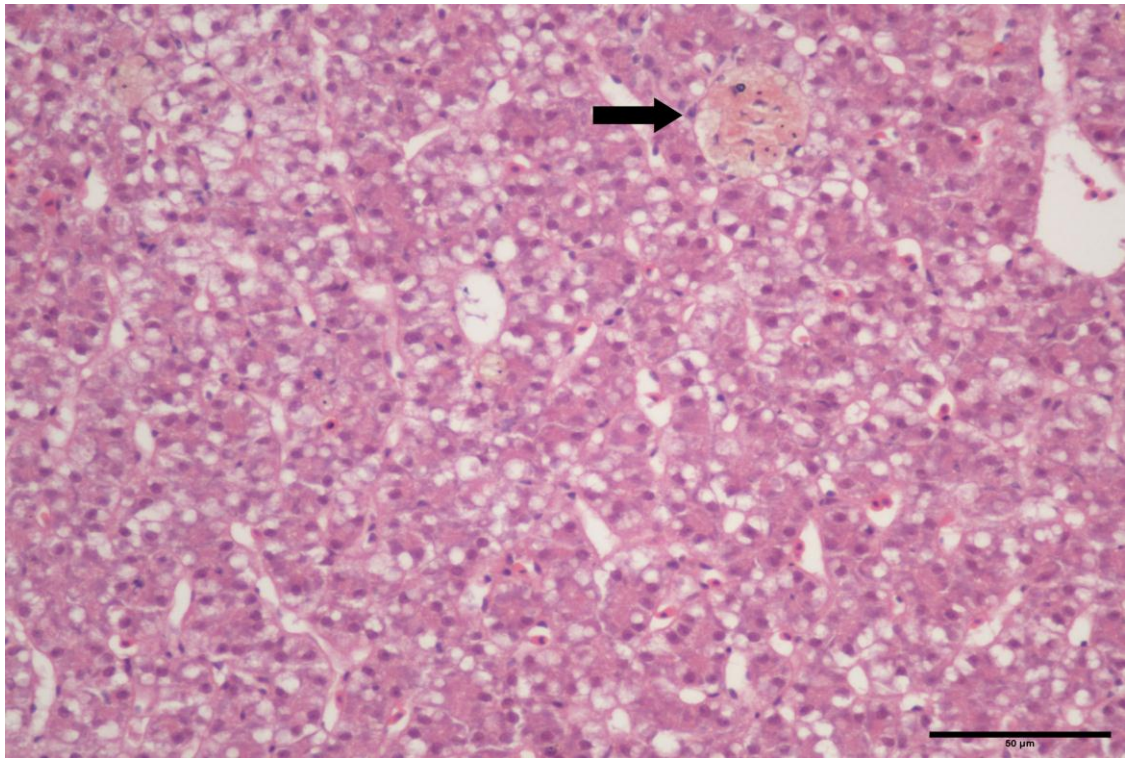
RA07045-3	H	South Cardigan Bay	Irish Sea	Female	2007	1, 2, 4
RA07045-541	H	North Cardigan Bay	Irish Sea	Female	2007	1, 2, 4
RA07045-539	H	North Cardigan Bay	Irish Sea	Female	2007	1, 2, 4
RA07045-452	H	St Bee's Head	Irish Sea	Female	2007	1, 2, 4
RA07045-455	H	St Bee's Head	Irish Sea	Female	2007	1, 2, 4
RA07045-459	H	St Bee's Head	Irish Sea	Female	2007	1, 2, 4

Table 2.1. Sample information. Cefas code, location of capture, sex and associated liver lesions of dab. Abbreviations; HCA: hepatocellular adenoma, ST: surrounding tissue around HCA, H: healthy (non-cancerous) dab liver, 1: bisulfite sequencing PCR (Chapter 4), 2: transcriptomic study (Chapter 4), 3: methylated-DNA immunoprecipitation coupled to high-throughput DNA sequencing (Chapter 4), 4: targeted metabolomics study (Chapter 5), 5: methylated-DNA immunoprecipitation coupled to cDNA flounder microarray (Chapter 4).

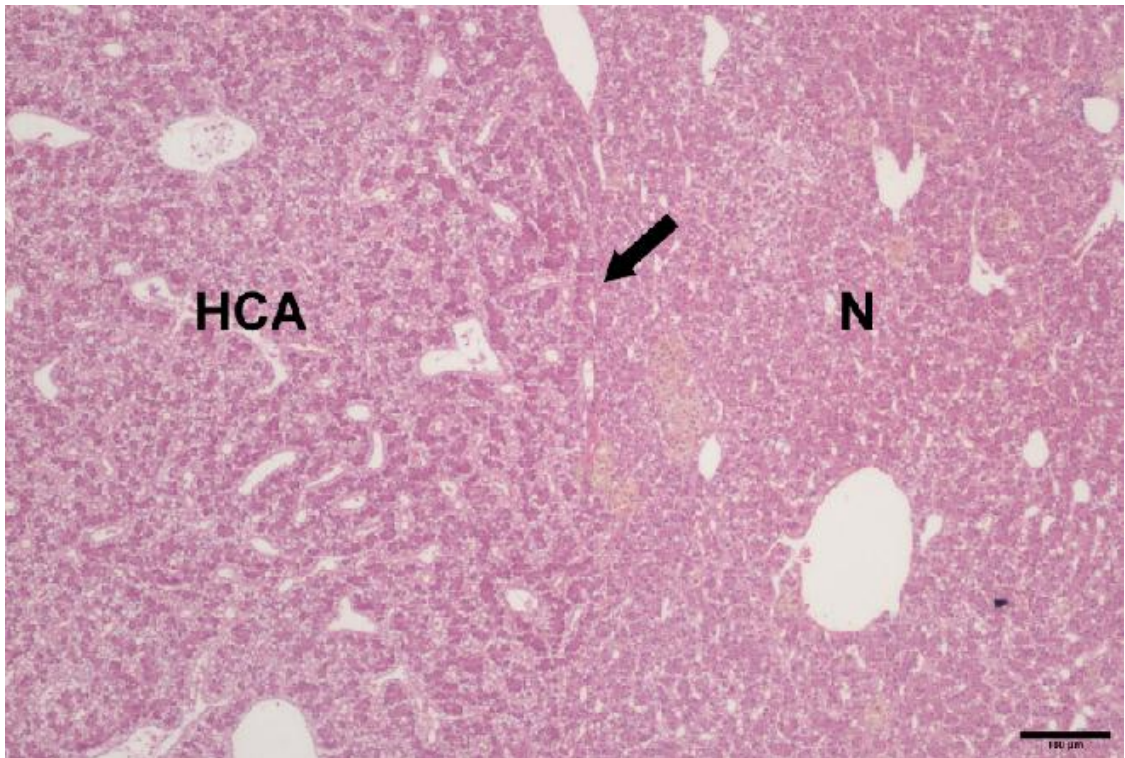
2.2.2.2. Histopathology

Cefas carried out all subsequent histopathology. Using standard histological protocols, fixed samples were processed to wax in a vacuum infiltration processor (Feist et al., 2004). A rotary microtome was used to prepare 3_5 µm tissue sections. The prepared sections were mounted onto glass slides for haematoxylin and eosin staining. Stained sections were analysed by light microscopy (Eclipse E800, Nikon, UK). The type of liver tumour was established based on guidelines described by Feist et al., (2004). Lesions were classified as HCA based on a combination of distinct morphological criteria, including presence of lesion with relative lack of macrophage aggregates, compression of surrounding parenchyma, thickened trabecular structure and the relative absence of atypical nuclear and cellular profiles (Feist et al., 2004) (Figure 2.2).

A.



B.



C.

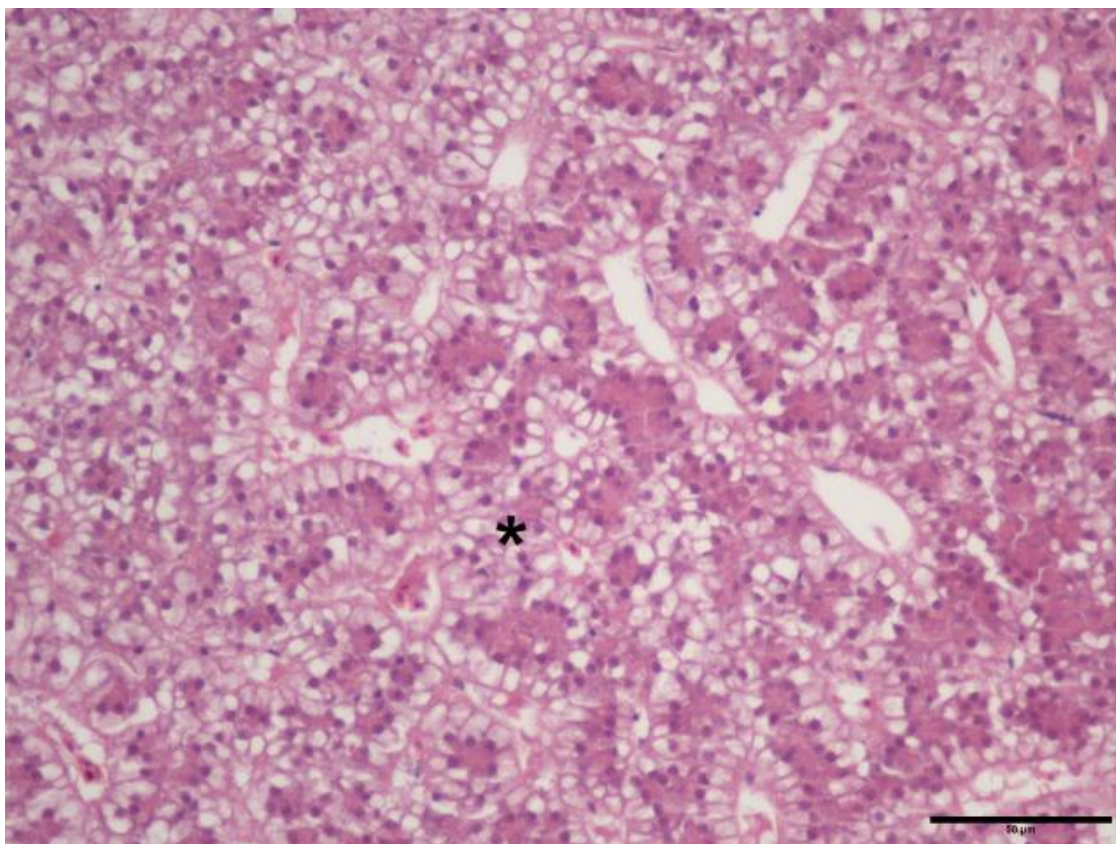


Figure 2.2. Confirmation of the type of dab liver lesion. Figure A: Normal liver architecture exhibiting hepatocytes cords of 2-3 cell thickness between individual blood vessels. The arrow shows the presence of macrophage aggregates. Figure B: Large uniform hepatocellular adenoma (HCA) exhibiting architectural differences and slight compression (arrow) against surrounding normal parenchyma (N). Macrophage aggregates can also be seen at periphery of lesion. Figure C: Abnormal cellular architecture within hepatocellular adenoma. Lesion has clear trabecular structure with thickening (*) of hepatocytes cords.

2.2.2.3. Sample preparation

Dab liver tissues provided by Cefas were kept at -80°C until extraction. Each liver was homogenised in 10 μ l/mg (wet tissue mass) water (Mass spectrometry grade, Fisher Scientific, Loughborough, UK) using a Precellys-24 ceramic bead-based homogeniser for two bursts of 10 seconds at 3800 x g (Stretton Scientific Ltd, UK). Aliquots of the homogenised tissues (no less than 10 mg) were used for RNA and DNA extraction. Separate aliquots of the

homogenised tissues (no less than 5 mg, with ideal amount of 10 mg) were used for metabolite extraction.

2.2.2.3.1. DNA and RNA extraction

To prevent degradation of the RNA samples by RNase, RNase free eppendorfs (Axygen, Inc, USA), water and barrier tips were utilised at all times.

Aliquots of the homogenised tissues (up to 30 mg of tissue) were added to ice cold RLT (600 µl) and β-mercaptoethanol (6 µl) mixture. Following centrifugation at full speed (16100 x g) for 3.5 minutes (Rotofix 32, Hettich Zentrifugen, Germany) and removal of the top lipid layer, RNA and DNA were extracted from the same homogenates using the AllPrep DNA/RNA kit (Qiagen Ltd, West Sussex, UK) according to manufacturers' guidelines. This kit allowed parallel isolation of DNA and RNA from the same sample using a membrane containing silica based resin with a particle size of 100 µm, a large pore size, and high density of positively charged diethylaminoethyl groups. Although both DNA and RNA samples are negatively charged, due to presence of phosphates of the DNA backbone, they differ on charge density and molecule size. DNA has a higher negative charge than RNA molecules; therefore, remains tightly bound to the diethylaminoethyl groups over a wide range of salt concentrations. RNA, protein, carbohydrates, and small metabolites are washed from Qiagen Resin with medium-salt buffers. DNA remains bound until eluted with a high-salt buffer. Salt residues are removed with ethanol washes, allowing the elution of the hydrophilic DNA samples in water based buffers.

The homogenate was transferred to Allprep DNA spin column and centrifuged (8000 x g, 30 seconds). The column was placed at 4°C for later extraction of DNA. To the supernatant containing the RNA sample, 70% (v/v) ethanol (350 µl) was added. Samples were mixed by pipetting and transferred to the RNeasy spin column placed in a 2 ml collecting tube and centrifuged (8000 x g, 15 seconds). Flow-through was discarded and buffer AW1 (700 µl) was added to the column and centrifuged (8000 x g, 15 seconds). Buffer RPE (500 µl) was added to the column and centrifuged (8000 x g, 15 seconds). Buffer RPE (500 µl) was used for washing the column for the second time and centrifuged (2 minutes, 8000 x g). The flow through and collecting tube was discarded and the column was placed in a clean collecting tube and centrifuged for an additional 1 minutes at full speed. The RNeasy spin column was

placed in a 1.5 ml RNase free eppendorf. RNase-free water was added directly to the spin column membrane and centrifuged (1 minute, 8000 x g) to elute the RNA.

The Allprep DNA spin column was removed from 4°C and buffer AW1 (500 µl) was added to the column and centrifuged (8000 x g, 15 seconds). Flow-through was discarded and buffer AW2 (500 µl) was added and centrifuged for 2 minutes at maximum speed. The Allprep spin column was placed in a clean 1.5 ml eppendorf. Care was taken to prevent any ethanol carryover. Elution buffer (100 µl) was pipetted directly onto the membrane of the column and left at room temperature for 1 minute followed by centrifugation (8000 x g, 1 minute) to elute the DNA. The quality of the RNA and DNA samples and the amount of the RNA and DNA samples were measured using a NanoDrop ND-1000 UV-VIS Spectrometer version 3.2.1 (Nanodrop, USA.). The A260/A280 ratio for all the samples ranged from 1.8 to 2.2. This indicated absence of protein contamination. RNA and DNA samples were stored at -80°C and -20°C, respectively until used in transcriptomic and DNA methylation studies.

2.2.2.3.2. Extraction of metabolites

Metabolites were extracted with a multi-sample ceramic bead-based system (Precellys 24, Stretton Scientific Ltd, UK) using a modified version of the protocol described by Southam et al., (2008). In the original method liver tissues were homogenised in a mixture of water and methanol. However, as in this study DNA, RNA and metabolites were extracted from the same sample and the methanol caused interference in RNA extraction, an alteration to the established method was made.

The extraction involved use of chloroform. To avoid any leaching of plasticisers from plastics into the extraction mixture via chloroform, any contact of chloroform with plastic was avoided.

Aliquots of the homogenised tissues prepared in section 2.2.2.3 (no less than 50 µl, equivalent to 5 mg of tissue, with ideal amount of 10 mg (100 µl)) were transferred to 1.8 ml glass vials containing an ice cold methanol: water mixture (40 µl:5.25 µl/mg wet tissue mass, HPLC grade methanol, Fisher Scientific) and vortexed (1 minute) at maximum speed. This was followed by the addition of ice-cold chloroform (40 µl/mg wet tissue mass, HPLC grade, Fisher Scientific, added using a Hamilton syringe, Fisher Scientific), water (20 µl/mg wet

tissue mass) and additional vortexing at maximum speed (30 seconds). Samples were placed on ice (10 minutes) to allow biphasic separation and then centrifuged at 1500 x g at 4°C for 10 minutes. Samples were left at room temperature for 5 minutes, leading to a clear separation into 3 layers. The upper layer containing water, methanol and polar metabolites, the lower phase containing chloroform and non-polar metabolites and lipids separated from each other via denatured proteins. Aliquots (50 µl) of the polar phase were transferred to eppendorfs using a glass Hamilton syringe and stored at -80°C. The non-polar phase was transferred to a glass vial and stored at -80°C. The non-polar layer was not used in the work presented in this thesis. Polar extracts were dried using a centrifugal evaporator (Thermo Savant, Holbrook, NY) and stored at -80°C until analysis.

2.2.3. European flounder (*Platichthys flesus*) liver tissue samples and calf (*Bos taurus*) thymus DNA

The four healthy flounder liver samples used in Chapter 3 were provided by Dr Tim Williams, School of Bioscience, the University of Birmingham. These liver samples were dissected from immature fish with unknown sex. DNA extraction procedure for these liver samples was similar to the methodology used for dab samples described in sections 2.2.2.3.1. DNA from calf (*Bos taurus*) thymus was purchased from Sigma-Aldrich Chemical company (UK).

2.3. Global measurement of DNA methylation levels

2.3.1. Enzyme-linked immunosorbent assay (ELISA)-based method

The Methylamp Global DNA Methylation Quantification kit (Epigentek Group Inc, Brooklyn, USA) was used according to the manufacture's guidelines for quantification of global DNA methylation. The methylated DNA was immobilised onto the surface of the wells and bound via 5-methylcytosine antibody and quantified using ELISA. This method is suitable for limited starting materials; however, the variability and accuracy of the method is lower than high-performance liquid chromatography (HPLC) and LC-tandem mass spectrometry (MS/MS).

Wells for samples, blank and control were placed in a plate frame included in the kit. GM2 solution (28 µl) was added to each well followed by addition of 2 µl of sample (1µg DNA/1

μl) and mixed. For blank, 30 μl of GM2 and for control 29 μl of GM2 and 1 μl of GM3 was added. Wells were placed in the oven at 37°C for 2 hours, and 60°C for 30 minutes for complete evaporation of solution. This was followed by the addition of GM4 (150 μl) and incubation at 37°C for 45 minutes. Wells were washed 3 times with 150 μl diluted GM1 (GM1 was diluted at 1:10 ratio with distilled water) and 50 μl of diluted GM5 (GM5 was diluted at 1:1000 ratio with diluted GM1) was added to each well and incubated at room temperature for 1 hour. Following incubation, wells were washed 4 times with diluted GM1 (150 μl). Diluted GM6 (50 μl, GM6 was diluted at 1:1000 ratio with diluted GM1) was added per well and incubated at room temperature for 30 minutes. Wells were washed 5 times with diluted GM1 (150 μl) and GM7 (100 μl) and incubated in the dark for 10 minutes. After 10 minutes, stopping buffer GM8 (50 μl) was added and the absorbance was read at 450 nm.

Percentage of methylation was calculated according to the following formula:

Percentage of methylation= OD (sample-blank) / OD (control-blank) x 100

2.3.2. HPLC and LC-MS/MS

The sample preparation procedure for both techniques is very similar. RNA contamination is eliminated from the DNA samples via enzymatic digestion of RNA. DNA samples are further enzymatically digested to form nucleotide monophosphates.

2.3.2.1. Removal of RNA contamination of DNA samples for both HPLC and LC-MS/MS

Using TE buffer (10 mM (hydroxymethyl)aminomethane (Tris)-HCl, pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA)) the total volume of the DNA sample (no less than 1 μg) was adjusted to 300 μl. RNase A (lyophilised RNase A was dissolved in 10 mM Tris, PH 7.5, 15 mM NaCl) was added to a final concentration of 100 μg/ml followed by the addition of RNase T1 to a final concentration of 2000 units/ml (lyophilised RNase T1 was dissolved in 20 mM Tris, pH 7.5). The solution was gently mixed and incubated at 37°C for 2 hours. After incubation, an equal volume of phenol/chloroform/isoamylalcohol (25/24/1, pH 8.0) was added to the sample, mixed and centrifuged at 16000 x g for 2 minutes. The aqueous layer, containing DNA, was transferred to a clean eppendorf and DNA was precipitated by the addition of 0.1 volume of 3M sodium acetate (pH 5.2), 2.5 volumes of absolute ethanol and

centrifugation at maximum speed for 10 minutes. The supernatant containing the hydrolysed RNA was discarded. The DNA pellet was washed with 70% (v/v) ethanol and centrifuged at maximum speed for 10 minutes. The supernatant was removed and the DNA pellet was air dried.

2.3.2.2. DNA hydrolysis prior to HPLC

Following removal of RNA contamination (2.3.2.1.), DNA was re-suspended in 100 µl deoxyribonuclease (DNase) 1 digestion buffer (10 mM Tris-HCl, pH 7.2, 4 mM magnesium chloride, 0.1 mM EDTA) and DNase1 was added to the final concentration of 50 µg/ml. Samples were incubated at 37°C for 16 hours. The non-specific endonuclease activity of DNase I results in cleavage of the phosphodiester bonds of both double- and single-stranded DNA. This results in formation of mono and oligonucleotides with 5'-phosphate and a 3'-hydroxyl groups. Following 16 hours of incubation, 2 volumes of 30 mM sodium acetate (pH 5.2) and zinc sulfate to a final concentration of 1 mM were added, as nuclease P1 is a zinc dependent nuclease. This was followed by the addition of nuclease P1 dissolved in nuclease P1 buffer (50 mM Tris HCl, pH 8.0, 10 mM EDTA) with a final concentration of 50 µg/ml and incubation at 37°C for 7 hours. This enzyme catalyses the non-specific endonucleolytic cleavage of single stranded DNA. This results in the formation of nucleoside 5'-phosphates. After incubation, samples were stored at -70°C or directly injected (50µl) into the HPLC.

2.3.2.3. DNA hydrolysis prior to LC-MS/MS

The DNA pellet prepared in section 2.3.2.1, following removal of RNA contamination, was re-suspended in sterile water (50 µl). To denature the DNA, DNA sample was heated (100°C, 3 minutes) and placed immediately on ice. A solution containing ammonium acetate (1M) and zinc chloride (1.25 µl; 45 mM, pH 5.1) was added to the sample. This was followed by the addition of nuclease P1 dissolved in nuclease P1 buffer (5 µl; 1 mg nuclease P1 was dissolved in 1 ml nuclease P1 buffer; nuclease P1 buffer: 50 mM Tris-HCl, pH 8.0, 10 mM EDTA) and incubation at 37°C for 2 hours. Following incubation, Tris (5 µl; 1.5 M, pH 8.0) and alkaline phosphatase (5 units; 0.2 units in 1 µl of 100 mM ammonium bicarbonate, pH 8.0) were added and the sample was incubated for 2 hours at 37°C. The hydrolysed DNA was filtered

using 0.2 µm filter (Millipore, Biomanufacturing and Life Science Research, UK) and syringe and stored at -20°C until analysed.

2.3.2.4. Standards

2'-deoxycytidine 5'-monophosphate ($C_9H_{12}N_3O_7P$) (dCMP), 5-methyl deoxycytidine 5'-monophosphate, disodium salt ($C_{10}H_{14}N_3O_7PNa_2$) (5mdCMP) (Reliable Biopharmaceutical Company, Missouri, USA), 2'-deoxythymidine 5'-monophosphate ($C_{10}H_{15}N_2O_8P$) (dTMP), 2'-deoxyguanosine 5' monophosphate ($C_{10}H_{14}N_5O_7P$) (dGMP), 2'-deoxyadenosine 5'-monophosphate ($C_{10}H_{14}N_5O_6P$) (dAMP), uridine 5'-monophosphate ($C_9H_{13}N_2O_9P$) (UMP), solutions (1 mg/ml) were prepared in DNase digestion buffer (2 parts nuclease P1 buffer, 1 part DNase 1 digestion buffer; DNase 1 digestion buffer: 0.01 M Tris-HCl, pH 7.2, 0.1 mM EDTA, 4 mM magnesium chloride). Retention times of each of the standards were established by injecting 50 µl of each of the prepared solutions. The separation of the bases was investigated by the injection of a mixture of the 5 bases. Contamination with RNA was monitored using a standard for UMP.

2.3.2.5. Reverse-phase high-performance liquid chromatography

All the solutions used for HPLC were prepared using HPLC grade solvents. The prepared solutions were filtered and de-gassed (Trans sonic T460, CamLab).

The method described by Ramsahoye (2002) was used for the measurement of global DNA methylation levels. A combination of factors, such as type of the column (solid phase), mobile phase, temperature and pH will affect the retention time and the observed order of separation of the components of a solution. The HPLC systems used were AKTA Explorer 10 with P900 pump and automatic UV detector (Amersham Biosciences) and Dionex P580 pump (Idstein, Germany) with the same column and guard column. The samples were injected onto a column containing a non-polar hydrocarbon chemically bonded to the surface of a rigid silica particles (solid phase) while an aqueous based mobile phase was running at a constant pressure and speed (1 ml/minute) through the column. The components of the injected sample are separated and observed in order of their affinity to the mobile phase. At the beginning of each run and prior to running the aqueous based mobile phase (50 mM diammonium orthophosphate, 50 mM orthophosphoric acid, pH 4.1), the column (APEX C18 column, 250

x 4.6 mm internal diameters, 5 µm particle size; Waters: HPLC Ltd, UK, Phenomenex, UK) was first washed for 30 minutes with 20% (v/v) methanol followed by HPLC grade water (30 minutes). After establishing a stable base line with the mobile phase, samples were injected (50 µl). At the end of each run the column was washed with water and 20% (v/v) methanol and stored in 20% (v/v) ethanol.

2.3.2.6. Nucleotide detection

Nucleotides were detected using a UV absorbance detector at the wavelength of 280 nm. This wavelength was used based on the λ_{max} of dCMP and 5mdCMP (dCMP: 272.7 nm, 5mdCMP: 278 nm). Unicorn 3.2 and Chromeleon 6.8 software (Dionex, Idstein, Germany) were used to calculate the peak area. Molar equivalents (ME) were calculated for 5mdCMP and dCMP using the peak areas and the extinction coefficients (dCMP: 9.3×10^3 , 5mdCMP: 11.8×10^3 , TMP: 10.2×10^3 , dGMP: 13.7×10^3 , dAMP: 15.3×10^3). This was followed by calculating the percentage of DNA methylation as shown below.

Molar equivalent (ME) = peak area/extinction coefficient ($\text{M}^{-1} \text{cm}^{-1}$)

% of DNA methylation = $(\text{ME}_{5\text{mdCMP}} / (\text{ME}_{5\text{mdCMP}} + \text{ME}_{\text{dCMP}})) \times 100$

2.3.2.7. LC-MS/MS

The prepared hydrolysed DNA samples were sent to the Department of Toxicology, University of Wurzburg, Germany for measurement of the percentage of methylated DNA via LC-MS/MS and by an independent group.

The hydrolysed DNA samples (1 µg) were diluted with mass spectrometry grade water (49.5 µl; Fisher Scientific, Loughborough, UK). The diluted samples (10 µl) were injected into the LC-MS/MS system. LC-MS/MS analysis was performed using an Agilent 1100 series liquid chromatography system coupled to an API 3000 triple quadrupole mass spectrometer equipped with Turbo ion spray source (Applied Biosystems, Darmstadt, Germany). Separation was achieved using a column (Reprosil Pur C18 column, 150 x 2 mm internal diameters, 5 µm particle size, Dr Maisch HPLC, Ammerbuch-Entringen, Germany) under a gradient running buffer (solvent A: 0.1% formic acid, solvent B: methanol) with the following conditions: 90%

A and 10% B (starting condition) followed by an increase to 40% B within the first 3 minutes and a linear increase to 100% B in 2.5 minutes with a flow rate of 300 μ l/minutes.

Analytes were detected in the positive ion mode with data acquisition performed as multiple reaction monitoring (MRM) of mass transitions of 268.2 mass to charge ratio (m/z) (parent compound) to 152.1 m/z (product) for 2-deoxyguanosine and mass transitions of 242.17 m/z (parent compound) to 108.95 m/z (product) for 5-methyldeoxycytidine. Quantification of 2-deoxyguanosine and 5-methyldeoxycytidine was performed by using a serial dilution of known amounts of external standards and generation of a standard curve for the two compounds of interest. The concentrations of 2-deoxyguanosine represented the total concentration of 2-deoxycytidine in the samples. Global DNA methylation was expressed as percentage of methylated cytidine in total sample.

2.4. Methylated-DNA immunoprecipitation (MeDIP)

Methylated fragments of DNA were separated from the un-methylated fragments using MagMeDIP kit (Diagenode, Belgium). The optimum starting amount of DNA for each MeDIP reaction is 1 μ g. The recommended starting amount of DNA prior to MeDIP for DNA methylation based microarray experiments is 5 μ g. Two options are commonly used by research groups: 1. Whole genome amplification of the MeDIP products to generate the sufficient starting material for DNA methylation microarrays. However, preferential amplification of fragments of DNA can bias the results. 2. MeDIP of five identical aliquots of the same DNA sample followed by combining the final immunoprecipitated DNA fragments from the five aliquots. As the latter approach remove the complications encountered during the first method, this approach was used in the present study.

DNA samples were dissolved in TE buffer (6 μ g in 300 μ l) and fragmented to generate 200 bp_1000 bp products using a sonicator (SONICS Vibra Cell, 100 watt, 3 x 10 seconds with 35 seconds intervals on ice with 20% amplitude). Lengths of the generated fragments were checked on a 1% agarose gel. The final concentration of the sonicated DNA was adjusted to 0.1 μ g/ μ l. A mastermix containing water (48 μ l), magbuffer A (24 μ l), magbuffer B (6 μ l) was prepared. Prepared mixture (65 μ l) was added to 10 μ l of DNA (0.1 μ g/ μ l, five aliquots of each sample), incubated at 95°C for 3 minutes and transferred immediately onto ice. In a separate 200 μ l tube, 137.5 μ l magbuffer A (1): water (5) mix was added to 11 μ l of

resuspended beads and centrifuged (5 minutes, 230 x g). Supernatant was discarded and beads were washed again with 137.5 µl magbuffer A (1): water (5), centrifuged (5 minutes, 230 x g) and supernatant was discarded. The washed beads were suspended in 22 µl of the magbuffer A (1): water (5) mix and kept on ice until used. In a second 200 µl tube the following mixture was prepared per IP reaction and was added to the heated DNA samples prepared earlier: antibody (1): water (1) mix (0.3 µl), magbuffer A (0.6 µl), water (2.10 µl), magbuffer C (2 µl). Following addition of the resuspend washed beads (20 µl); samples were placed on a rotating wheel overnight at 4°C.

The beads were washed after the overnight incubation. All washing steps were performed in the cold room (4°C). IP samples were removed from the rotator, briefly centrifuged in a bench top centrifuge in the cold room and placed on a magnetic rack for 1 minute. Supernatant was removed, ice cold magwash buffer 1 (100 µl) was added and the beads were resuspended by inverting the tube, and then samples were incubated on the rotating wheel (4 minutes, 4°C). Following incubation, samples were briefly centrifuged, placed on the magnetic rack for 1 minute and supernatant was discarded. Beads were washed 3 times with magwash buffer 1 and once with magwash buffer 2, following the above procedure. Eppendorfs containing the washed bead pellets were placed on ice and the remaining work was conducted at room temperature with the samples kept on ice. A mixture containing buffer D (360 µl), buffer E (40 µl) and buffer F (16 µl) was added to the pellets prior incubation in a thermo-shaker (10 minutes, 65°C, 230 x g). One volume of phenol/chloroform/isoamyl alcohol (25:24:1) was added to the samples. Samples were incubated at room temperature for 10 minutes on a rotating wheel prior to centrifugation (2 minutes, room temperature 230 x g). The aqueous phase was transferred to a clean eppendorf and 1 volume of chloroform/isoamyl alcohol (24:1) was added to the samples and centrifuged (2 minutes, room temperature, 230 x g). The aqueous layer was transferred to a clean eppendorf and a mixture containing glycogen (0.02 mg, Fermentas Life Sciences), co-precipitant (5 µl), precipitant (40 µl) and absolute ethanol (1 ml) was added to the samples, vortexed and incubated (-80°C, 90 minutes). Following incubation, samples were centrifuged (30 minutes, 4°C, 230 x g). Ethanol (70% v/v, 500 µl) was added to the supernatant prior to centrifugation for 20 minutes at 230 x g at 4 °C. Supernatant was removed, DNA pellets were air dried and stored at -80°C until subsequent use.

2.5. Microarray experiments

2.5.1. Pilot study: cDNA flounder microarray

The flounder cDNA microarray was constructed for European flounder (*Platichthys flesus*) as part of Genomic Tools for Biomonitoring of Pollutant Coastal Impact (GENIPOL) project in 2006. The procedure for construction of this microarray is described in detail in Williams et al., (2006). Briefly, 13,824 PCR amplicon clones derived from flounder cDNA, relating to approximately 3336 genes of European flounder (*Platichthys flesus*) were spotted onto a microarray slide (Corning UltraGAPS, Corning Inc., Corning, NY, USA) by an MGII robot (Biorobotics, Cambridge, UK) employing a 48 split pin tool. Each clone was spotted twice for a total of 27,648 distinct elements per microarray, including control DNA elements. Clones had been sequenced and annotated by BLAST against the GenBank database (www.blast.ncbi.nlm.nih.gov/Blast.cgi). As the library was redundant, many genes were represented by several overlapping clones; therefore, the intensity reading for all the overlapping probes representing a gene was averaged prior to data analysis. Sequences that were from the same gene that shared no overlapping regions were retained as separate identities and could either represent different genetic isoforms or different regions of the same gene.

As sequence homology is sufficiently great between the two closely related flatfish, European flounder and dab, the use of this microarray for transcriptional studies in dab is valid (Diab et al, 2008; Cohen et al., 2007; Small et al., 2010). Hence, as a pilot study this microarray was employed in combination with MeDIP for investigation of DNA methylation alterations between dab HCA samples and surrounding tissues. Samples used in this experiment are indicated with number 5 in Table 2.1 section 2.2.2.1. An overview of the MeDIP-flounder cDNA microarray is given in Figure 2.3.

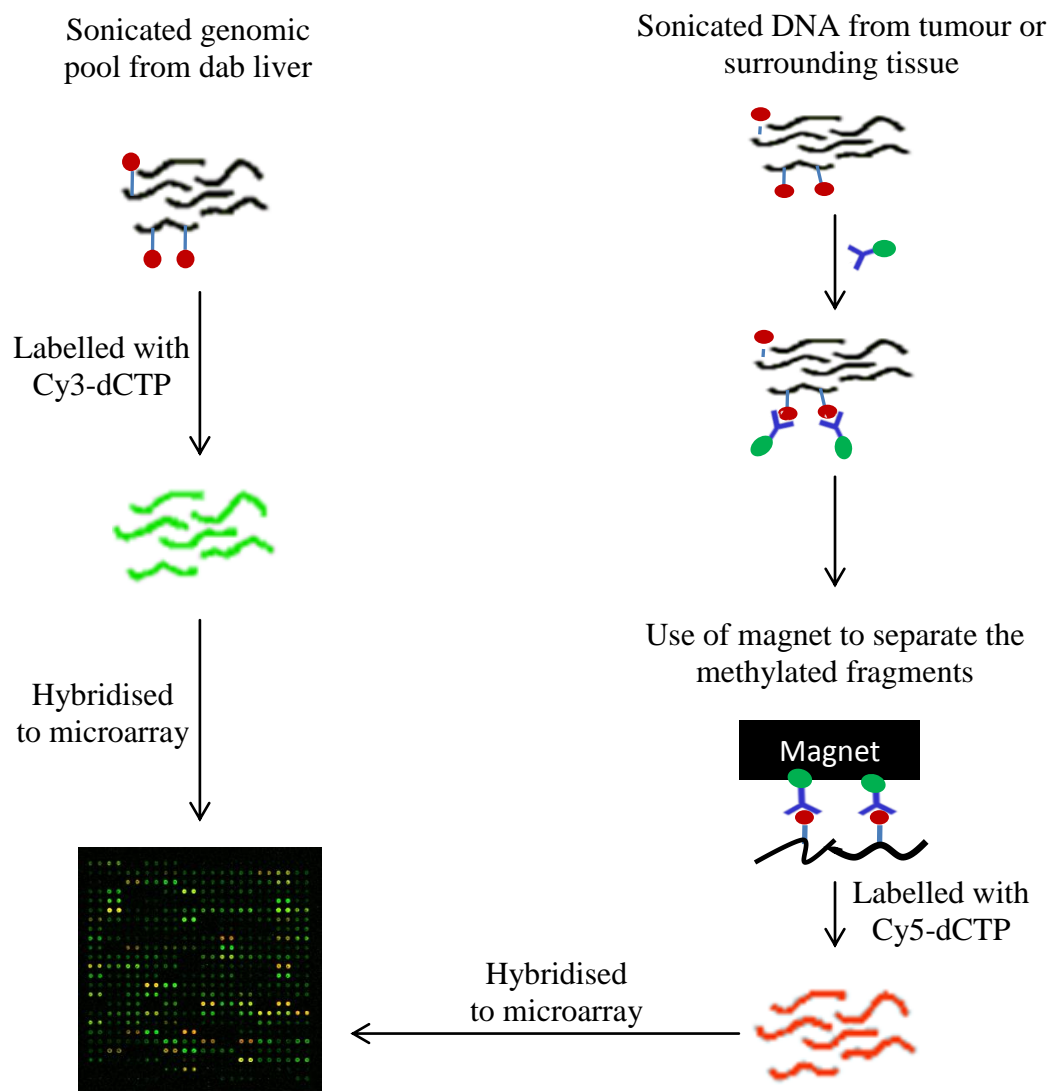




Figure 2.3. Overview of the dab MeDIP-flounder cDNA microarray experiment. DNA from dab hepatocellular adenoma (HCA) samples and surrounding liver tissues (ST) were sonicated, immunoprecipitated using an antibody against 5-methylcytosine and labelled with cyanine-5 labelled deoxycytidine triphosphate (Cy5-dCTP). Sonicated DNA samples from both groups of analysis were labelled with cyanine-3 labelled deoxycytidine triphosphate without immunoprecipitation (Cy3-dCTP). Cy3-dCTP and Cy5-dCTP labelled samples were hybridised to a flounder cDNA microarray. : methylated cytosine, : 5-methylcytosine antibody attached to a magnetic bead.

2.5.1.1. DNA labelling and purification

Sonicated, MeDIPed DNA fragments from each biological sample were labelled with Cy5-dCTP and the reference DNA sample was labelled with Cy3-dCTP. As these dyes are light sensitive care was taken to minimise their exposure to light. 2.5xRandom priming mix (20 µl) (Bioprime DNA labelling kit, Invitrogen, Paisley, UK) was added to the DNA samples and the final volume was adjusted to 41 µl using sterile water. To denature the DNA, samples were heated (100°C, 10 minutes) and placed immediately on ice. This was followed by the addition of 10xdNTPs (5 µl), each at 1.2 mM except dCTP at 0.6 mM (Bioline Ltd., UK), Cy3-dCTP or Cy5-dCTP (1 µl; GE Healthcare, Amersham, UK) and Klenow polymerase (1 µl, equivalent to 40 units; Invitrogen) to the samples while they were on ice. This was followed by incubation of the samples in the dark (37°C, 18 hours). Following 18 hours of incubation, labelled DNA samples were purified and eluted in Tris-Cl buffer (10 mM, pH 8.5, 60 µl; also known as elution buffer) using QIAquick spin columns (Qiagen Ltd, West Sussex, UK) according to the manufacturer's protocol. Briefly, 0.1 volume of sodium acetate (3M, pH5.2) and 5 volume of PB buffer were added to 1 volume of labelled genomic DNA samples. Samples were mixed and transferred to QIAquick spin columns and centrifuged (1 minute, 5900 x g). Flow through was discarded and buffer PE (750 µl) was added to the columns and centrifuged for 1 minute at 5900 x g. After discarding the flow through, columns were centrifuged for 1 minute. Columns were placed in 1.5ml eppendorfs and water (60 µl) was pipetted directly onto the membrane of the column. To elute the purified labelled DNA samples, columns were centrifuged for 1 minute.

The concentration of labelled DNA samples were measured using a spectrophotometer (for Cy3-dCTP labelled samples, an intensity reading of 1 at A550 equated to 0.13 pmoles/µl of dye incorporated; for Cy5-dCTP labelled samples, an intensity reading of 1 at A650 equated to 0.08 pmoles/µl of dye incorporated). Cy5-dCTP labelled test DNA (60 pmoles) was mixed with labelled genomic pool (60 pmoles). The volume was adjusted to 30 µl with water or, if required, concentrated using microcon 30 excision filter (Millipore, Biomanufacturing and Life Science Research, UK).

2.5.1.2. Hybridisation

For hybridisation 30 μ l of 2x hybridisation buffer (50% formamide, 10x sodium citrate sodium chloride (SSC), 0.2% sodium dodecyl sulphate (SDS), 1.6 μ g/ μ l poly-A (Alta Bioscience, Birmingham, UK) and 4 μ g/ μ l yeast tRNA (Sigma-Aldrich Chemical Company, UK)) (Sambrook et al., 1989) was combined with 30 μ l of the labelled mixture. Samples were heated (2 minutes at 95°C) and centrifuged (30 seconds) followed by pipetting of the mixture under a 25 \times 60L coverslip (VWR, Lutterworth, UK) placed on a pre-hybridised slide. The pre-hybridised slide was located in a hybridisation chamber containing water (20 μ l) and incubated overnight at 42°C. To prepare the microarray slides for hybridisation, slides were placed in a pre-hybridisation solution (25% formamide, 5x SSC, 0.1% SDS, 10 mg/ml bovine serum albumin (BSA)) for 2-10 hours at 42°C. Slides were washed twice with water and once with absolute ethanol and dried by centrifugation at 400 x g for 10 minutes (Hettich Rotofix 32, Hettich Labtechnology, Tuttlingen, Germany).

2.5.1.3. Washing and scanning

After incubation slides were washed in clean 50 ml centrifuge tubes (Falcon, Becton Dickinson Labware, Le Pont De Claix, France) on a Denley Spiramix 5 (Thermo Scientific, Langenselbold, Germany) with 2 \times SSC, 0.1% SDS solution for 5 minutes at 42°C, followed by 0.1 \times SSC, 0.1% SDS solution for 5 minutes at room temperature, and four washes of 0.1 \times SSC solution for 5 minutes each. Slides were quickly dipped in two water washes, one absolute ethanol wash and dried by centrifugation (10 minutes, 400 x g). Slides were placed in a slide container and protected from exposure to light.

Fluorescent signals were detected using an Axon 4000B laser scanner and Genepix software (Molecular Devices, Sunnyvale, CA, USA) (Figure 2.4). Spots were checked manually and spots with poor morphology were excluded from further analyses and labelled as 'not found'. Spot intensity was taken as the median value of all pixels in the spot, minus the median value of pixels surrounding the spot.

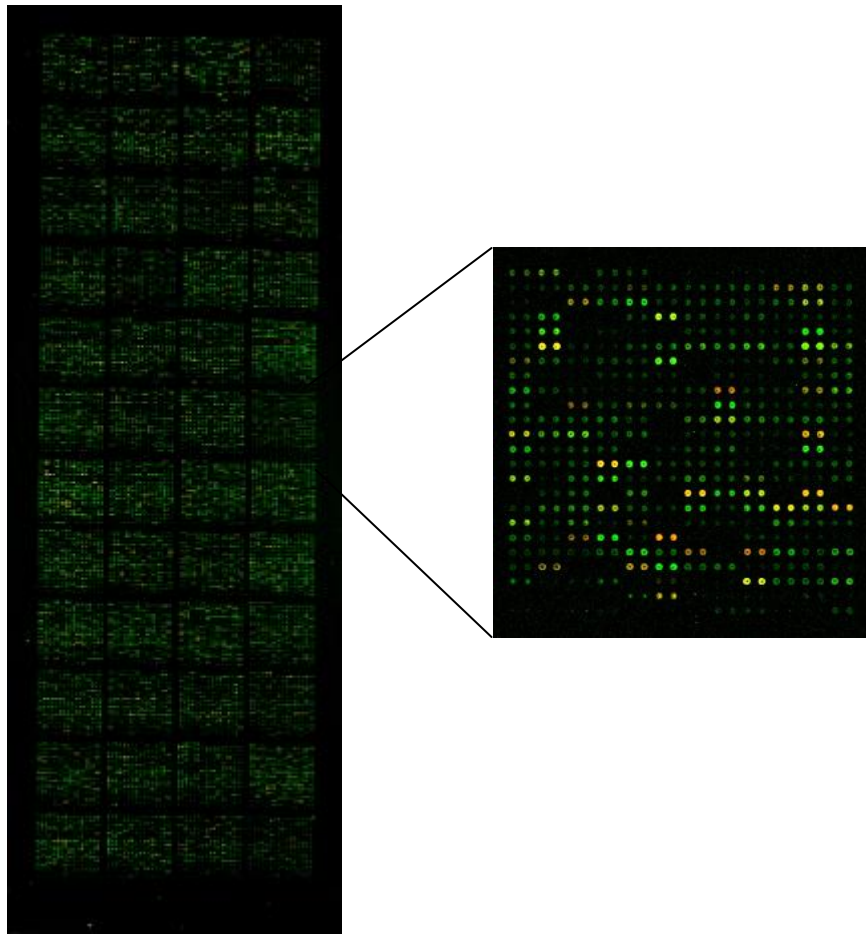


Figure 2.4. Scanned image of a flounder cDNA microarray from MeDIP-cDNA experiment. The microarray contains a total of 27,648 elements printed on 48 sub-arrays. Each sub-array contains 576 spots and each clone is spotted in duplicate.

2.5.2. CGI (1.5kb downstream to 1kb upstream of TSS) zebrafish tiling microarray experiment

The experimental procedure used is similar to the methodology described in section 2.5.1 for MeDIP-cDNA flounder microarray experiment (Figure 2.3). In this experiment 4 male zebrafish HCC samples and 4 male zebrafish liver control samples were used.

2.5.2.1. Design of the 4x44k format CGI (1.5kb downstream to 1kb upstream of TSS) zebrafish tiling microarray

Probes were designed by Genotypic (Genotypic Technology, Bangalore, India) and printed by Agilent (Agilent technologies, Berkshire, UK). The criteria described by Gardiner-Garden and Frommer (1987) for prediction of CpG islands in vertebrate were used to predict CpG islands in zebrafish, with minor modifications. These criteria were: percentage of C+G $\geq 50\%$, length ≥ 100 bp and observed/expected ratio ≥ 0.6 . The zebrafish genome sequence derived from Ensembl database (version 56, genome build zv8, genome build date April 2009) was employed. Repetitive sequences were masked and the Emboss CpG prediction tool (<http://www.ebi.ac.uk/Tools/emboss/cpgplot/>) identified 9,192 genes containing 14,507 CpG islands in the region of 1.5kb upstream to 1kb downstream of transcription start sites (TSS). The 5' end of each transcript was used as TSS. From the preliminary list of genes (9,192), 60-mer probes were designed with an average spacing of 25 bp where possible. Due to probe specificity, probes could not be constructed for all predicted CGIs and this also influenced the probe spacing. Probes with multiple BLAST hits on the zebrafish genome and with an identical 55 bp alignment were removed. Finally, 43,960 probes were designed for 7,903 CGIs on 6,024 genes. Slides were printed containing 45,220 features; 1,227 Agilent control features, 33 blank features and 43,960 probes. The microarray design was named Agilent Birmingham *D.rerio* 025794 45220v1 and its design is available from ArrayExpress under accession A-MEXP-1813. Probes were functionally annotated against the zebrafish genome and additional gene ontology (GO) terms were found using Blast2GO (Gotz et al., 2008).

2.5.2.2. DNA labelling

The Agilent Genomic DNA enzymatic labelling kit (Agilent technologies, Berkshire, UK) and the protocol provided were used with minor modifications for labelling the MeDIPed test DNA and the genomic pool. Immunoprecipitated DNA was dissolved in nuclease free water (26 μ l), random primers (5 μ l) were added and samples were mixed by pipetting. This was followed by incubation at 95°C for 3 minutes and 5 minutes incubation on ice. A labelling mastermix containing 5xbuffer (10 μ l), cyanine-3 labelled deoxyuridine triphosphate (Cy3-dUTP) or cyanine-5 labelled deoxyuridine triphosphate (Cy5-dUTP) (1mM, 3 μ l), Exo-Klenow fragment (1 μ l) was prepared on ice and added to the sample. To clean the labelled

samples, 1xTE buffer pH 8.0 (430µl) was added to the samples and transferred to a microcon 30 size exclusion filter (Millipore, Biomanufacturing and Life Science Research, UK) and centrifuged at 8000 x g for 10 minutes. After discarding the flow-through 1xTE buffer pH 8.0 (480 µl) was added and the samples were re-centrifuged for concentrating the samples to a final volume of 13 µl. Qualities of the labelled samples were checked using NanoDrop spectrophotometer. Samples with a Cy5 pmol/µg yield of >7 and Cy3 pmol/µg yield of >18 were used. Equal pmol (80 pmol) of Cy3-dUTP and Cy5-dUTP labelled DNA samples were mixed and the final volume was adjusted to 22.5 µl using RNase/DNase free distilled water.

2.5.2.3. Hybridisation

Sample hybridisation was performed using Agilent Oligo aCGH/Chip on chip hybridisation kit. 10x blocking agent was prepared by addition of nuclease free water (1250 µl). The following components were added in order to a 1.5 ml eppendorf: Cy3-dUTP/Cy5-dUTP labelled DNA mixture (22.5 µl), Cot-1 DNA (5 µl; Invitrogen), Agilent 10x blocking solution (11 µl), de-ionized formamide (16.5 µl) and Agilent 2x hybridization buffer. The prepared mixture was centrifuged, heated (3 minutes, 95°C) and transferred to a circulating water bath for 30 minutes. After incubation samples were centrifuged (1 minute, room temperature, 17900 x g). Samples (100 µl) were loaded into individual wells of a gasket slide held in the base of a hybridisation chamber (Agilent technologies, Berkshire, UK). A 4x44k format microarray was placed on top of the gasket slide. Chambers were reassembled and clamped and placed in an Agilent hybridisation oven (24 hours, 67°C, 3 x g) (Figure 2.5).

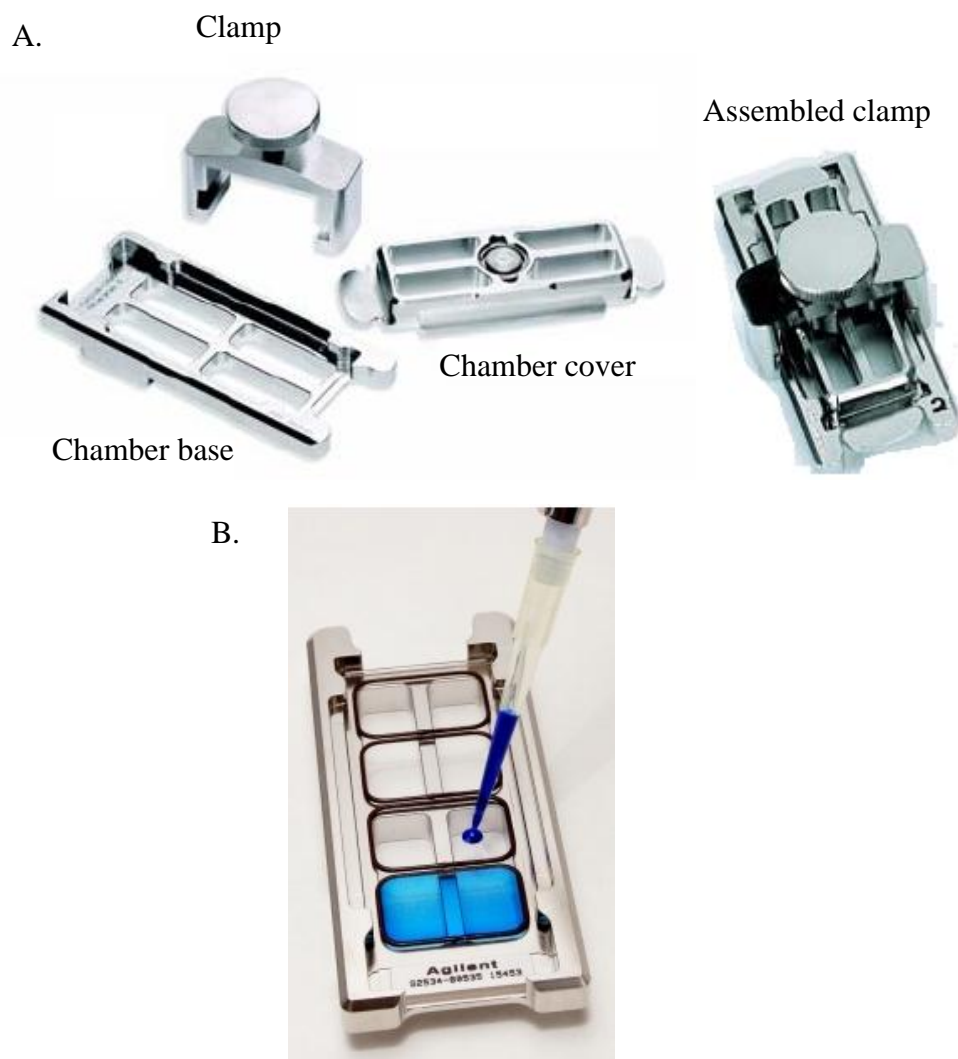


Figure 2.5. Agilent chamber. A. The individual components and an assembled Agilent hybridisation chamber. B. To load the samples gasket was placed in the chamber base and each sample was pipetted into the 4 individual wells of the gasket slide. The microarray was placed on top of the gasket slide. The chamber cover was placed on top of the microarray slide and the chamber base and cover were clamped securely together.

2.5.2.4. Washing and scanning

Following overnight hybridisation, slides were separated from the gasket slides and washed using oligo aCGH wash buffer 1 and 2, stabilisation and drying solution (Agilent technologies, Berkshire, UK). Slides were dipped in a slide staining dish containing wash buffer 1 at room temperature to remove the gasket. Next, they were placed in a second dish

containing the wash buffer 1. After placing the slides, the solution was stirred for 5 minutes using a magnetic stirrer. After 5 minutes, slides were transferred to a staining dish containing wash buffer 2 (37°C, 1 minute) while stirring with a magnetic bar. Slides were immediately dipped in acetonitrile followed by dipping in stabilisation and drying solution. The dried slides were scanned using an Axon 4000B laser scanner and Genepix software (Molecular Devices, Sunnyvale, CA, USA) (Figure 2.6).

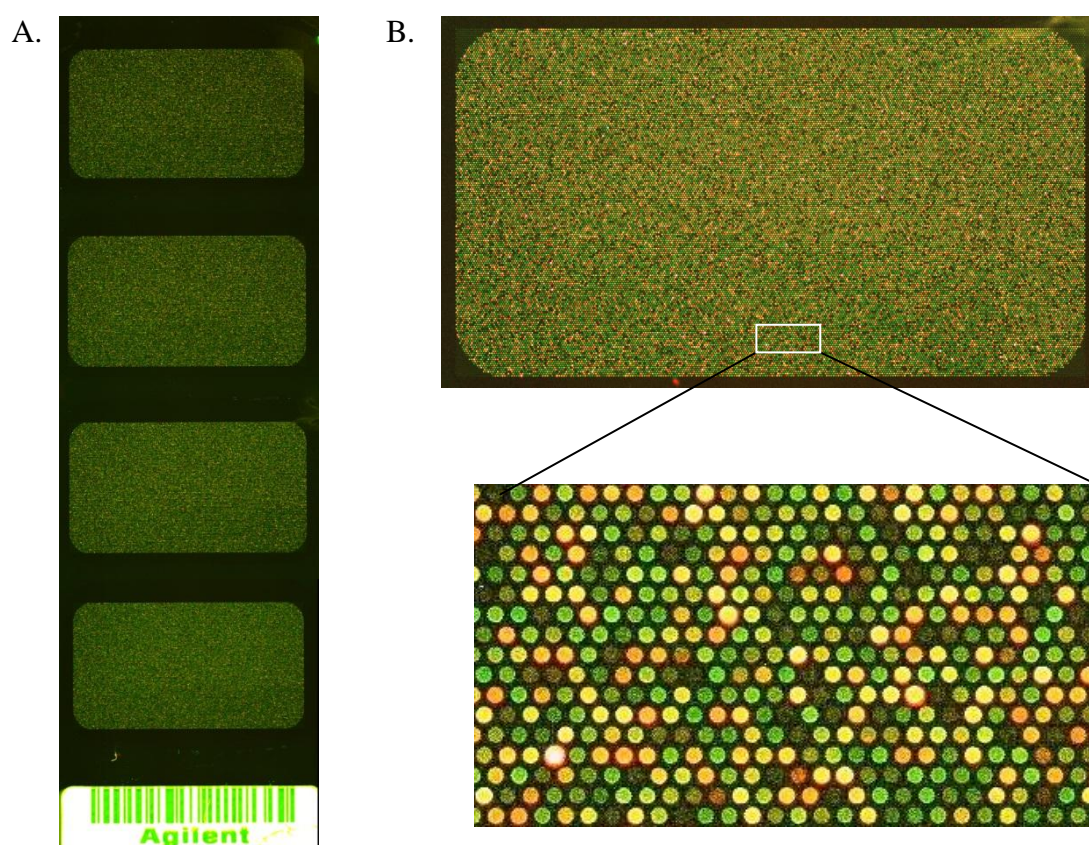


Figure 2.6. Zebrafish 4x44k tiling microarray. A. Scanned image of a zebrafish 4x44k tiling microarray. Each of the four sections contains 43,960 probes and act as an individual microarray. B. Enlarged image of a sub-array. The green colour represents lower signal intensity from the test sample compared to the reference sample. The red colour represents higher signal intensity from the test sample compared to the reference sample. The yellow colour indicates equal signal intensity from both samples.

2.5.3. Dab MeDIP *de novo* high-throughput sequencing and transcriptomic profiling

DNA extracted from a hepatocellular adenoma sample and corresponding surrounding tissue, as described in section 2.2.2.3.1, were subjected to immunoprecipitation using 5-methylcytosine antibody as described in section 2.4. Samples used for this experiment are indicated in Table 2.1 section 2.2.2.1. The immunoprecipitated fragments of DNA were subject to *de novo* DNA high-throughput sequencing (HTS) using an Illumina Genome Analyser II (Illumina, Inc, USA). This was conducted at Beijing Genomic Institute (BGI), as briefly described in section 2.5.3.1. The data achieved from the MeDIP-HTS aided the design of a dab specific gene expression microarray. The samples used in this experiment are shown in Table 2.1, section 2.2.2.1 and the experimental procedure is explained in section 2.5.3.3. An overview of the described experimental procedures is shown in Figure 2.7.

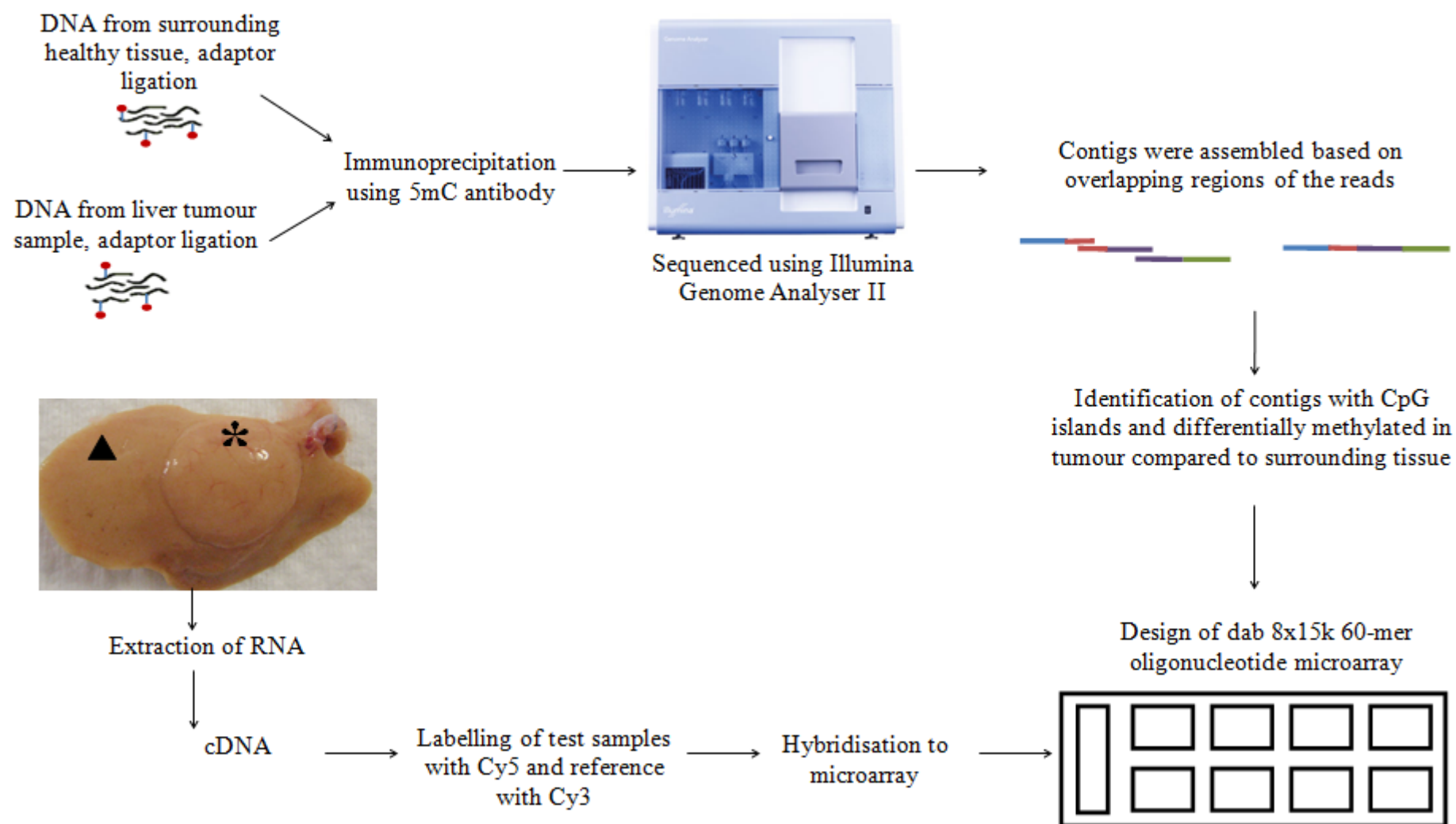


Figure 2.7. Overview of the MeDIP, HTS, assembly of contigs, design of the dab microarray and microarray experimental procedure.

* Dab HCA tumour, ▲ surrounding tissue

2.5.3.1. MeDIP *de novo* high-throughput sequencing

The MeDIP *de novo* high-throughput sequencing was conducted at Beijing Genomics Institute (BGI, China) based on methods described by Li et al., (2010). Briefly, DNA samples extracted from one HCA sample and corresponding surrounding tissue (5 µg) were sonicated. Fragmentation was carried out using a Covaris sonication system (16 cycles of duty cycle: 10%, intensity: 5, cycles/burst: 200, time: 16 minutes; Covaris, Inc, USA) which resulted in 100-500 bp rich DNA fragments. The Illumina Paired-end library preparation kit (Illumina, Inc, USA) and the protocol provided were used to generate adaptor-attached DNA fragments for sequencing. During this process the overhangs from fragmentation were converted into blunt ends using T4 DNA polymerase and Klenow polymerase. "A" base was added to the 3' end of the blunt end phosphorylated DNA fragments which provided a recognition site for a single "T" base overhang containing adaptors. Methylated adaptor ligated DNA fragments were separated from the un-methylated fragments using the methodology described in section 2.4.

The methylated fractions were purified and eluted in Tris-Cl buffer (10 mM, pH 8.5, 27 µl) using ZYMO DNA Clean and Concentrator-5 kit (Zymo Research Corporation, CA, USA), following the manufacturer's instructions. Briefly, 2_7 volumes of DNA binding buffer was added to each volume of sample, mixed and transferred to Zymo-Spin column. The column was centrifuged (30 seconds, 10000 x g) and the flow through was discarded. Column was washed with wash buffer (200 µl), centrifuged (30 seconds, 10000 x g) and the flow through was discarded. The column was transferred to a clean eppendorf and the sample was eluted by direct addition of Tris-Cl buffer (10 mM, pH 8.5, 27 µl) to the membrane of the column and centrifugation at 10000 x g for 30 seconds. The eluted DNA was amplified by adaptor-mediated PCR in a final reaction volume of 50 µl. The reaction contained purified DNA (23 µl), Phusion DNA polymerase mix (25 µl) and PCR primers (2 µl). Cycling parameters consisted of: 94°C for 30 seconds, 16 cycles of 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 30 seconds with a final step of 1 minute at 72°C. Products were sequenced using Illumina Genome Analyser II (Illumina, Inc, USA). Raw data for both samples, hepatocellular adenoma HCA-18-RA09065-830 and healthy surrounding tissue ST-19-RA09065-830, were submitted to NCBI database with accession numbers GSE31124 and GSM770685, respectively.

2.5.3.2. Identification and annotation of differentially methylated regions

HTS of MeDIP genomic DNA from the HCA sample and ST sample produced 89,925,735 and 87,826,470 paired-end reads, respectively. The SOAP *de novo* software (a member of Short Oligonucleotide Analysis Package family) was used to assemble the reads with a 25 k-mer overlap to achieve sets of contiguous sequences (contigs) for both samples. The assembled contigs later contributed to microarray design as described in section 2.5.3.3.

In order to identify differentially methylated regions that contained CGIs in the HCA sample compared to ST, a third assembly was generated by combining the reads from both individual samples (HCA and ST). Contigs were generated using the list of combined reads from both samples and SOAP *de novo* software. N50 contigs with length <200bp were discarded and the remaining N50 contigs were defined as “methylated DNA fragments”. (Note: N50 contig size was defined as the length of the smallest contig S in the sorted list of all contigs where the cumulative length from the largest contig to contig S was at least 50% of the total length). Reads from HCA and ST were aligned to the “methylated DNA fragments” assembled by SOAP software. The sequences that were uniquely mapped with ≤ 2 bp mismatches were retained. This assembly resulted in 264,008 contigs with lengths ≥ 200 bases, of which 69,046 passed the CGI criteria (observed/expected ≥ 0.6 and CG% $\geq 50\%$), with an average length of 323.22 bases. As the numbers of short reads contributed by each sample to each contig was known, the ratio of enrichment between HCA and ST was calculated to identify candidate with differentially methylated CGIs. The identified candidate CGIs (1,693 contigs) were annotated using BLASTx versus Genbank non-redundant (nr) protein database employing Blast2GO at E-value $< 1E-06$ (Gotz et al., 2008; Conesa et al., 2005).

2.5.3.3. Design of flatfish dab specific 8x15k oligo microarray based on contigs achieved from MeDIP *de novo* HTS

The oligonucleotide microarray was designed via Agilent eArray (Agilent technologies, Berkshire, UK). The microarray was designed using the 1,693 differentially methylated contigs with known protein-coding sequences. The 1,693 contigs were achieved from the HTS run, as described in section 2.5.3.1. However, these contigs were supplemented with additional data to produce a more comprehensive resource for dab transcriptomics. Eight four of the most highly differentially methylated unidentifiable CGIs from the HTS results were

included. This necessitated the design of both forward and reverse microarray probes for these specific CGIs. Additional putative protein coding sequences were derived from the CGI contigs that displayed no differential DNA methylation (7416), the HCA and ST assemblies >60 bases that aligned with flatfish expressed sequence tags (ESTs: 1599) and that matched with other protein coding sequences (3864) and 5 gene sequences from other flatfish species. These were manually curated to reduce duplication of probes for sequences aligning with the same orthologs. In total 14,951 sequences were uploaded to Agilent eArray, resulting in the successful design of 14,919 60-mer probes. The microarrays were printed in 8x15k format with each sub-array including 14,919 experimental spots, 290 empty positions and 77 negative and 459 positive Agilent control spots (Agilent-031032_8x15k_BhamDab). The microarray details and experimental design (Transcription profiling of dab hepatocellular adenoma liver) are available from ArrayExpress under accessions A-MEXP-2084 and E-MTAB-734, respectively.

2.5.3.4. Quality assessment

In addition to using NanoDrop ND-1000 UV-VIS Spectrometer version 3.2.1 (Nanodrop, USA) for the assessment of the quality of RNA samples as described (section 2.2.2.3.1), the integrity of the RNA samples and the assessment of contamination of the RNA samples with DNA, was evaluated using 2100 Bioanalyzer and a Eukaryote Total RNA Nano assay chip and Agilent small RNA kit (Agilent Technologies, Berkshire, UK). Briefly, Gel Matrix (550 µl) was transferred to the filter columns provided with the kit and centrifuged (10 minutes, 1500 x g). An aliquot of the filtered Gel Matrix (65 µl) was transferred into an eppendorf and the remaining Gel Matrix was stored at 4°C for later use. Dye (1 µl) was added to the Gel Matrix, vortexed and centrifuged at 13000 x g for 10 minutes. The prepared mixture was placed on ice in the dark. A chip was placed in the chip priming station. Avoiding any debris and air bubbles, the gel-dye mix (9 µl) was pipetted into the well of the chip marked as G with a black dot. Plunger was set to 1 ml mark followed by closing of the chip priming station. The plunger was pressed down until it reached the chip surface and was held in this position for 30 seconds and then released. After an additional 5 seconds the plunger was slowly pulled back up to the 1 ml mark and the chip was removed from the chip priming station.

An additional gel-dye mix (9 μ l) was pipetted into the two wells of the chip marked as G with grey shading and the remaining mixture was discarded. NanoMarker buffer (5 μ l) was pipetted into each of the 12 sample wells and the ladder well. To eliminate any secondary structures, the RNA samples and the ladder were denatured by heating (2 minutes, 70°C). The heated samples and the heated ladder were briefly centrifuged and 1 μ l was loaded into the appropriate wells. The chip was vortexed using Agilent Chip Vortexer for 1 minute at 240 revolutions per minute (rpm) and placed in the 2100 Bioanalyzer. Equipment used are shown in Figure 2.8 and examples of Bioanalyzer trace is shown in Figure 2.9.

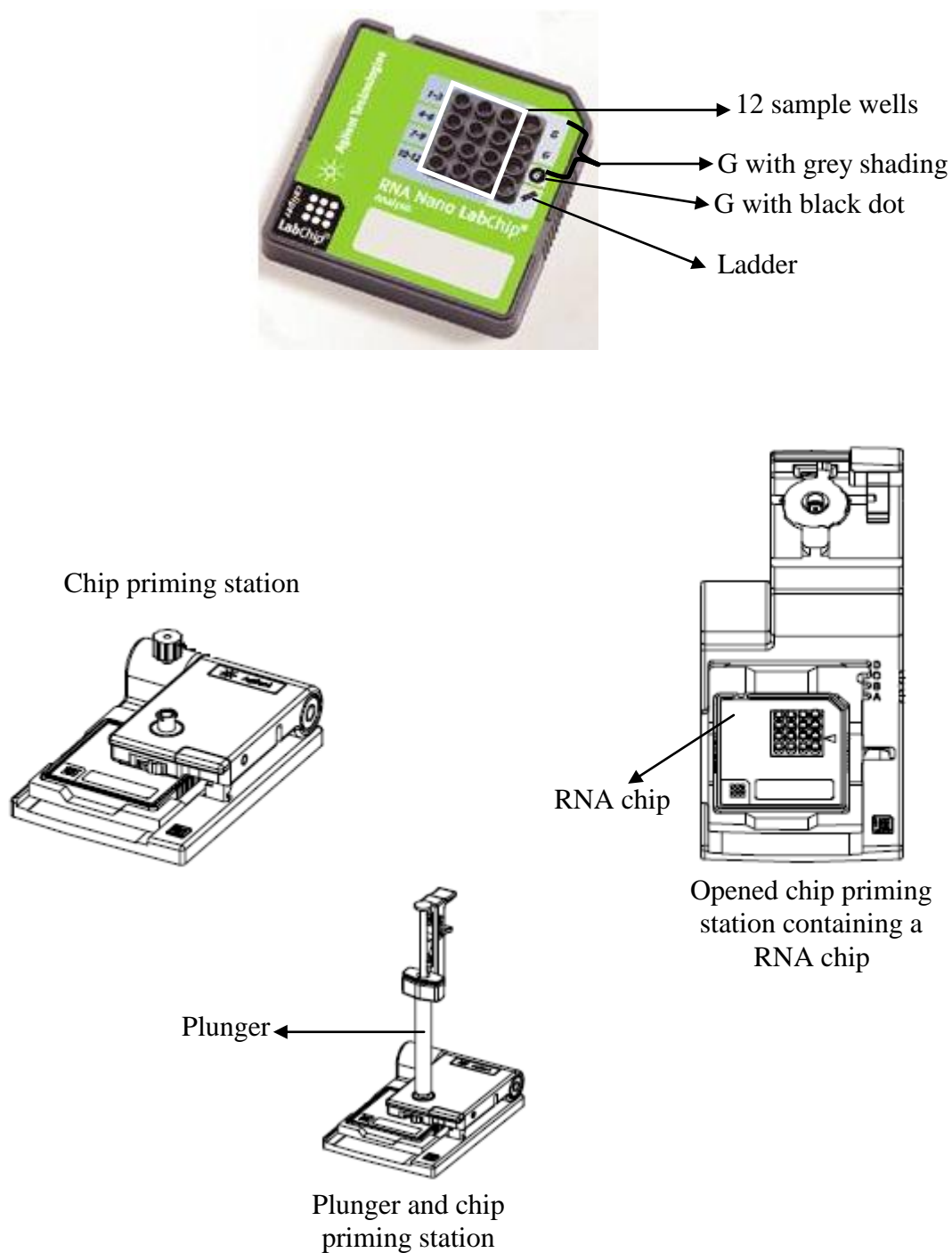


Figure 2.8. Equipment required for assessment of RNA integrity with Agilent 2100 Bioanalyzer. RNA chip with different sections of the chip, chip priming station, the location of the chip in the priming station and an assembled chip, priming station and plunger are shown in this figure.

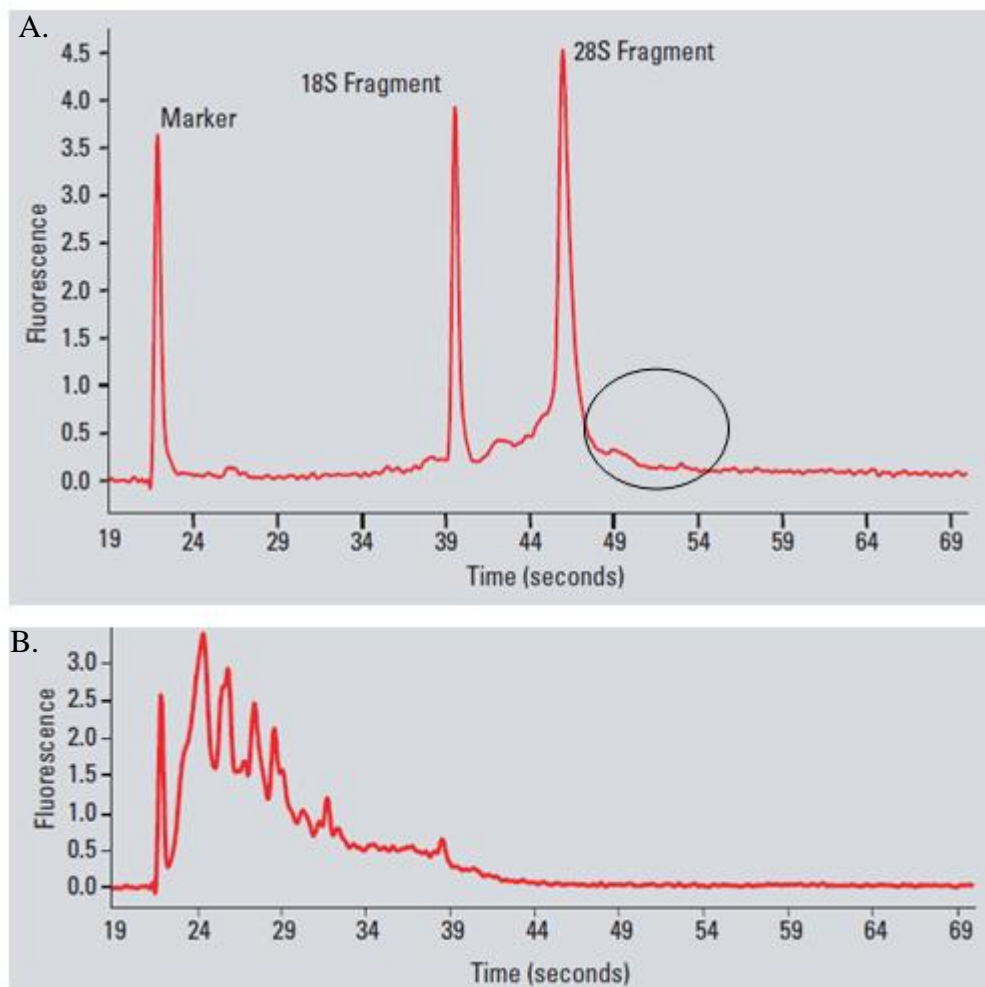


Figure 2.9. 2100 Bioanalyzer traces. A. An electropherogram example of a good quality RNA sample. Bioanalyzer marker: this is a 25 bp marker to standardise each of the lanes to the ladder that is run with each chip. The majority of the total RNA is 18S (2 kilobase pairs (kb)) and 28S (5 kb) ribosomal RNA. mRNA samples comprise a small percentage of the total RNA ($\sim <3\%$) and therefore, it is difficult to assess their quality directly. Thus rRNA quality is used as an indicator of the quality of the underlying mRNA population. Clear bands for 18S and 28S rRNA is indicative of intact RNA sample. Contamination of RNA samples with DNA is shown as a peak in the region indicated with a circle. B. Example of electropherogram of a degraded RNA sample. Images were adapted from Agilent website.

2.5.3.5. DNase treatment of RNA samples

RNA samples with genomic DNA contamination, detected by using the 2100 Bioanalyzer, were treated with genomic DNA-free kit (Ambion Applied Biosystems, Warrington, UK) in order to digest and remove any traces of genomic DNA.

Water was used to adjust the volume of the RNA samples (5 µg) to 25 µl. DNase I (0.5 µl) and DNase I buffer (2.5 µl) were added to the samples followed by incubation at 37°C for 30 minutes. After completion of the incubation time, DNase inactivation reagent was vortexed and 0.1 volume was added to the samples (with the maximum amount of 2 µl). This was followed by incubation of the samples at room temperature for 2 minutes, during which samples were mixed at least 3 times. To separate the digested DNA from RNA samples, samples were centrifuged (90 seconds, 10000 x g) and the supernatant containing the RNA sample was transferred to an RNase free eppendorf.

2.5.3.6. cDNA and cRNA synthesis and labelling

The Agilent Two Colour Low Input Linear Amplification Kit with Spike-ins (Agilent technologies, Berkshire, UK) and the protocol provided, were used for labelling. The total RNA from both HCA and ST samples were converted to cDNA, labelled with cyanine-5 labelled cytidine triphosphate (Cy5-CTP) and converted to complementary RNA (cRNA). The reference sample was generated by combining RNA from all samples. The total RNA from the reference sample was labelled with cyanine-3 labelled cytidine triphosphate (Cy3-CTP) and subsequently converted to cRNA.

Briefly, total RNA (50 ng) dissolved in water (1.5 µl) was supplemented with Agilent two-colour spike mix, T7 promoter primer (0.8 µl) and the volume was adjusted to 3.5 µl with nuclease free water. The sample was denatured by incubating the mixture at 65°C for 10 minutes followed by incubation on ice for 5 minutes. During this incubation time, the 5x first strand buffer was heated (4 minutes, 80°C) and a mastermix containing: 0.1 M dithiothreitol (DTT, 2 µl), 10 mM dNTP mix (0.5 µl), 5x heated First strand buffer (6 µl) and AffinityScript RNase block mix (1.2 µl) was prepared on ice. The prepared mastermix was added to the sample, mixed and incubated at 40° for 2 hours, 70°C for 15 minutes and 5 minutes on ice for the synthesis of cDNA.

The synthesised cDNA was subsequently labelled and converted to cRNA by addition of a mastermix containing nuclease free water (0.75 μ l), 5x Transcription buffer (3.2 μ l), 0.1 M DTT (0.6 μ l), NTP mix (1 μ l), T7 RNA Polymerase Blend (0.21 μ l) and Cy3-CTP or Cy5-CTP (0.24 μ l) depending on the type of sample. Following gentle mixing, samples were incubated at 40°C for 2 hours.

Any uncoupled Cy3-CTP and Cy5-CTP were removed using the RNeasy mini kit (Qiagen Ltd, West Sussex, UK) and the protocol provided. Briefly, nuclease free water (84 μ l), buffer RT (350 μ l) and absolute ethanol (250 μ l) were added to the labelled samples and mixed by pipetting. Samples were transferred to RNeasy columns and centrifuged (30 seconds, 15700 x g). Columns were transferred to clean collecting tubes and washed twice with buffer RPE (500 μ l) and centrifuged (30 seconds, 15700 x g). Columns were placed in RNase free eppendorfs and RNase free water (30 μ l) was added directly to the membranes of the columns. Samples were incubated at room temperature for 1 minute followed by centrifugation (30 seconds, 15700 x g) for eluting the cyanine labelled cRNA samples.

Labelling efficiency was determined by NanoDrop ND-1000 UV-VIS Spectrophotometer, version 3.2.1 (Nanodrop, USA). A yield of 0.825 μ g cRNA, and a specific activity (Cy3 and Cy5 dye incorporation) of >6 pmol Cy3/ μ g cRNA was considered sufficient.

2.5.3.7. Hybridisation

Microarray hybridisation was performed using an Agilent Gene Expression Hybridisation kit (Agilent technologies, Berkshire, UK) and according to the manufacturer's protocol. 10x blocking agent was prepared by the addition of nuclease free water (1250 μ l). The following components were added in order to a 1.5 ml eppendorf: Cy3 labelled cRNA (300 ng), Cy5 labelled cRNA (300 ng), 10x blocking agent (5 μ l), water (to adjust the volume to 24 μ l) and 25x fragmentation buffer (1 μ l). Samples were gently mixed and fragmented by incubation at 60°C for 30 minutes. The fragmentation reaction was stopped by incubation of the samples on ice for 1 minute. 2x GEx hybridisation buffer HI-RPM (25 μ l) was added to each sample followed by careful pipetting and centrifugation (1 minute, room temperature, 15700 x g). Samples were placed on ice and randomly loaded onto 8x15K format slides (Agilent-031032_8x15k_BhamDab) and hybridised overnight (65°C, 10 rpm).

2.5.3.8. Washing and scanning

Washing steps were similar to the methods described in section 2.5.2.4. The dried slides were scanned using an Agilent G2565CA microarray scanner system (Agilent technologies, Berkshire, UK).

2.6. Bisulfite sequencing PCR

2.6.1. Overview of the bisulfite sequencing PCR

Bisulfite sequencing PCR (BSP) is the ‘gold standard’ methodology for detection of 5-methylcytosine (5mC). This method provides a detailed analysis of the methylation status of individual CpGs within the amplified regions. During prolonged incubation of DNA with sodium bisulfite, cytosines in single stranded DNA are deaminated and converted to uracil while 5mC is immune to this transformation (Suzuki and Bird, 2008). This allows the distinction of methylated cytosines from remaining un-methylated cytosine. The sodium bisulfite treated DNA is amplified using BSP specific primers and then sequenced (Figure 2.10). CpH (H=A, C, T) regions within a sequence are used as internal controls for the measurement of efficiency of the sodium bisulfite treatment. Successful sodium bisulfite treatment will result in conversion of all cytosines located at these regions to thymine.

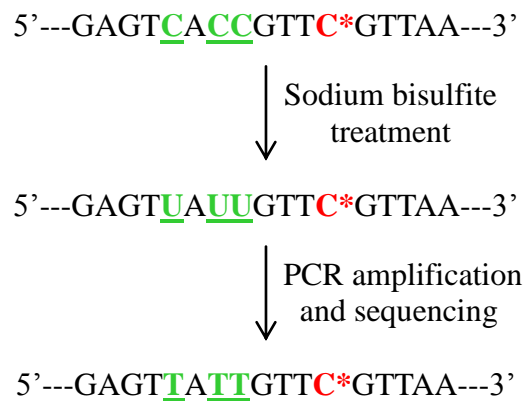


Figure 2.10. Bisulfite sequencing PCR. During bisulfite treatment, un-methylated cytosine is deaminated and converted to uracil. This allows the distinction of methylated cytosines from un-methylated cytosines. Green C: un-methylated cytosines, Red C with a star: methylated cytosines.

2.6.2. Sodium bisulfite treatment

The EZ DNA Methylation kit (Cambridge Biosciences, UK) was used for bisulfite conversion, according to the manufacturer's protocol. Prior to the start of the experiment, the amount of the DNA samples were adjusted to 400 ng and the CT-conversion reagent was dissolved in water (750 μ l) and M-dilution buffer (210 μ l), and then vortexed for 10 minutes at room temperature. The CT-conversion reagent is light sensitive; therefore, care was taken to minimise its exposure to light.

M-dilution buffer (5 μ l) was added to the DNA samples (400 ng) and the final volumes of the samples were adjusted to 50 μ l with water. Samples were incubated at 37°C for 15 minutes. After incubation, CT-conversion reagent (100 μ l) was added to each sample and the samples were incubated at 50°C for 16 hours in the dark. Following incubation, samples were placed on ice for 10 minutes. During this time M-binding buffer (400 μ l) was added to each Zymo-spin column placed in collecting tubes. After 10 minutes, samples were transferred to the columns and mixed with the M-binding buffer by inverting the columns several times and then the samples were centrifuged ($>10000 \times g$, 30 seconds). The flow through was discarded, M-wash buffer (100 μ l) was added to the columns and then the columns were centrifuged for 30 seconds at full speed. M-desulphonation buffer (200 μ l) was added to the columns and samples were incubated at room temperature for 20 minutes, and then centrifuged at full speed for 30 seconds. Columns were washed twice with M-wash buffer (200 μ l) and centrifuged twice for 30 seconds. Following the final wash, columns were placed in clean eppendorfs and M-elution buffer (10 μ l) was added directly onto the membranes, and then columns were centrifuged for 30 seconds to elute the bisulfite treated DNA samples.

2.6.3. Design of the bisulfite sequencing PCR primers

The MethPrimer software was used for identification of CpG islands (CGI) and design of BSP primers within a given DNA sequence (Li and Dahiya, 2002) (Figure 2.11). The criteria for prediction of CpG islands (CGI) in vertebrates (length ≥ 200 bp, observed/expected ≥ 0.6 , CG% ≥ 50), as described by Gardiner-Garden and Frommer (1987) were used for the identification of CGIs. The bisulfite treated DNA was used as a template for the design of the BSP primers. The primers were designed at either ends of the CGIs at locations free of CpG dinuclotides. This allowed un-biased amplification of both methylated and un-methylated DNA templates.

BSP primers only contain three nucleotide types. One primer contains A, G and T nucleotides and has an excess of T nucleotide while the other contains A, C and T nucleotides and has an excess of A nucleotide. The primers used in this thesis are shown in Tables 2.2 and 2.3.

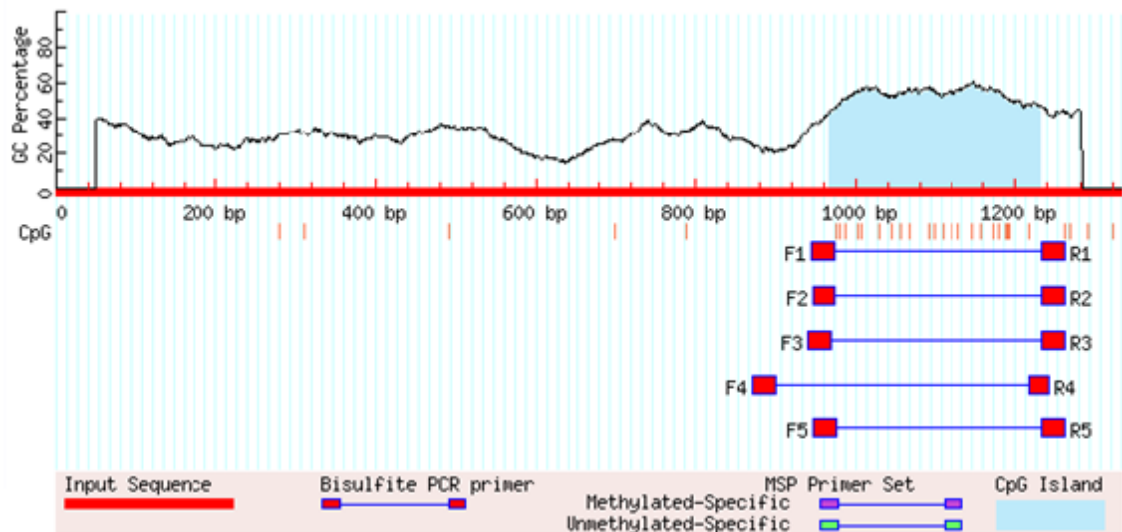


Figure 2.11. The methprimer programme used for identification of CGIs and design of bisulfite sequencing PCR primers. As an example, the sequence for the promoter region of the *coronin, actin binding protein 2ba* gene (*coro2ba*; ENSDARG00000079440) on chromosome 7 was imported into the methprimer software. The programme was set for design of BSP primers around CGIs. The region highlighted in light blue represents a CGI and the location of each CpG is shown with a red line.

Symbol	Chr.no	Gene ID	Primer	Annealing temp (°C)	Product size (bp)
<i>gstp1</i>	14	ENSDARG00000005039	F: 5'-TTATATTTATAGTTTTGGTTGTTTTT-3' R: 5'-CTACTCCTAATCTACAATACACTATC-3'	53.3	230
<i>Mycb-1</i>	2	ENSDARG00000007241	F: 5'-TAGAGTTGTTTGTTGGGTTTGTATT-3' R: 5'-AAAAC TACTTATCATAATATTCTTCC-3'	53.3	166
<i>Mycb-2</i>	2	ENSDARG00000007241	F: 5'-TTTTTTTATTGTTATATTTATGTTTTTATA-3' R: 5'-ACTACTTATCATTTACTACAAAATTTC-3'	53.3	226
<i>cdh1</i>	7	ENSDARG00000024371	F: 5'-ATTTATATTTGGAAATTGTGATTAT-3' R: 5'-AATTATTACACAAACCATTTA ACTC-3'	51.3	210
<i>cdk8</i>	24	ENSDARG00000016496	F: 5'-TTTTTAAATAAAGGGTTAAAATTTA-3' R: 5'-AATACAAAATATTATTAAACTAACAATC-3'	52.3	196
<i>timp2</i>	3	ENSDARG00000037806	F: 5'-TAAAATAGATGATGGATATGATGATTAAG-3' R: 5'-CCAACACTAACACAAATACAATAAAAAAC-3'	52.3	242
<i>hsp90ab1</i>	11	ENSDARG00000029150	F: 5'-TGTTTAAAGTTGAAAATAAATTTATAAT-3' R: 5'-ACTACTTTCAACCATTCACTACACC-3'	52.3	236
<i>mapk1</i>	5	ENSDARG00000027552	F: 5'-GAAAAGTTAATTGTTTTTATGGTAT-3' R: 5'-ATTAACAAACATATATTTTCTACTAATAAT-3'	52.3	183

Continued from previous page

<i>tp53</i>	5	ENSDARG00000035559	F: 5'-GATTTGGATGTAGAAAGTTTATTGGATA-3' R: 5'-CCCCTAATTACAAAAAAACATATAAC-3'	54.3	244
<i>msh6</i>	13	ENSDARG00000011666	F: 5'-TTTTTATTGGTTTATATGTAAATGA-3' R: 5'-CTCAAAACTACCCTACAATTATAAC-3'	53.3	118
<i>msh2</i>	12	ENSDARG00000018022	F: 5'-AAGTTTATAATAAGATAGTAATGTTTGT-3' R: 5'-TCCATTAATCTTAAAAACCTCTTTAACTAC-3'	53.3	285
<i>ntla</i>	19	ENSDARG00000009905	F: 5'-GGAAATATGTTTGTTTTAAGTTT-3' R: 5'-ATAACAATCATTTCATTAATAAACTCTTTA-3'	54.3	179
<i>dnmt4</i>	23	ENSDARG00000036791	F: 5'-ATTATTATTTTGAGAAAGGGTTAAATAGT-3' R: 5'-AAAAATAATCCACTACCAAAATACC-3'	52.3	230
<i>rap2c</i>	14	ENSDARG00000015649	F: 5'-TTGATTATTTAAGAGAGAGAAAATT-3' R: 5'-CAAAAAATTTATTAACCACACAAAC-3'	52.3	222
<i>cx43</i>	20	ENSDARG00000041799	F: 5'-TGGTTTAGTTTTTTTATTTATTAGGGTT-3' R: 5'-TAAATTTCCCTCCACAATCCAAAA-3'	53.3	197
<i>coro2ba</i>	7	ENSDARG00000079440	F: 5'-TGTTGTTATGTGTTTTGTAGATGTTATG-3' R: 5'-ACCTATATTTTACCATTAATTTAATATAAA-3'	53.8	315

Continued from previous page

<i>Igfbp1b</i>	2	ENSDARG00000038666	F: 5'-TAAGTGTAGTTTAAAAATATAAGTGTGAAA-3' R: 5'-TATACTCTAACAAAATAACCTCAAC-3'	53.8	271
<i>angptl3</i>	6	ENSDARG00000044365	F: 5'-TGTATAAATTTTTTGAGATTTTGGT-3' R: 5'-ATAACTAACTTATCCACAACAAAC-3'	53.8	180

Table 2.2. Bisulfite sequencing PCR primers used in zebrafish study (Chapter 3). *Glutathione S-transferase P1* (*gstp1*), *Transcriptional regulator Myc-B* (*mycb*), *cadherin 1* (*cdh1*), *Cell division protein kinase 8* (*cdk8*), *TIMP metalloproteinase inhibitor 2* (*timp2*), *Heat shock protein HSP 90-beta* (*hsp90ab1*), *Mitogen-activated protein kinase 1* (*mapk1*), *Cellular tumor antigen p53* (Tumor suppressor p53) (*tp53*), *MutS homolog 6* (*msh6*), *MutS homolog 2* (*msh2*), *No-tail gene* (*ntla*), *DNA (cytosine-5-)-methyltransferase 4* (*dnmt4*), *RAP2C, member of RAS oncogene family* (*rap2c*), *Gap junction alpha-1 protein (Connexin-43)* (*cx43*), *coronin*, *actin binding protein 2ba* (*coro2ba*), *insulin-like growth factor binding protein 1b* (*igfbp1b*) and *angiopoietin-like 3* (*angptl3*). F: forward primer, R: reverse primer.

Gene symbol	Contig ID	Primer	Annealing temp (°C)	Product size (bp)
<i>pcdh1g22</i>	6580896	F: 5'-GGGTTATTTGGTTATTAAGGTGA-3' R: 5'-TCTATAAACAACTCTAAAAAAATC-3'	53.3	183
<i>map1aa</i>	6176098	F: 5'-TTTGTGATTTGAATAAAGAGTTGTA-3' R: 5'-TCTCAAAAAACAAATACAAATAAATACTA-3'	52.4	165
<i>ttl7</i>	6540261	F: 5'-TATAGTTTTATATTTAAAATTTGGATTTTT-3' R: 5'-ATACTTATCCTTTATTTATTCTACCACTAA-3'	53.3	283
<i>nid1</i>	6612046	F: 5'-TGTTGAGTGGTTTTTAATTTGTTGT-3' R: 5'-AAATATTTATATATTCTCTCCTTCC-3'	52.4	224
non-coding protein region	6429668	F': 5'-GTTTGTGGGATTTATTGTGGATTA-3' R: 5'-CACAAAATTTAAATACTTCTACTATC-3'	53.3	154

Table 2.3. Bisulfite sequencing PCR primers used in dab study (Chapter 4). *protocadherin 1 gamma 22* (*pcdh1g22*), *microtubule-associated protein 1aa* (*map1aa*), *tubulin tyrosine ligase-like member 7* (*ttl7*), *nidogen 1* (*nid1*). F: forward primer, R: reverse primer.

2.6.4. Amplification of the sodium bisulfite treated DNA

CpG islands are highly repetitive regions. In addition, treatment with sodium bisulfite further increases the receptiveness of these regions by increasing the percentage of T and A bases. Hence, specific amplification conditions and DNA polymerase enzyme are required for amplification of the sodium bisulfite treated DNA samples. Zymo *Taq* DNA polymerase (Cambridge Biosciences, UK) was used for all the work on bisulfite sequencing PCR presented in this thesis.

A mastermix containing 2x reaction buffer (25 µl), dNTP mix (0.5 µl; each dNTP at 0.25 mM), bisulfite treated DNA (3 µl) as template, Zymo *Taq* DNA polymerase (0.4 µl; 2 units/50 µl), forward primer (10 pmol) and reverse primer (10 pmol) was prepared. The total volume was adjusted to 50 µl with water. The samples were amplified using a thermocycler (Minicycler™, MJ Research, UK) with the following programme: 95°C for 10 minutes, 38 cycles of 95°C for 30 seconds, variable annealing temperatures for 35 seconds and 72°C for 45 seconds. Following completion of the cycling phase, a final extension step was performed at 72°C for 7 minutes.

Each primer set was optimised and validated prior to use in the experiments. The products were checked by running the PCR products on agarose gel. The products were subsequently sequenced and compared with the expected sequences.

2.6.5. Generation of artificially methylated and un-methylated genomic DNA

Genomic DNA can be artificially methylated via treatment with CpG methyltransferase, *Sss* I, in the presence of *S*-adenosylmethionine (SAM). Following bisulfite treatment, amplification of the gene of interest, and sequencing, all CpG dinucleotides should be 100% methylated. In addition, un-methylated genomic DNA can be generated by conducting whole genome amplification. During amplification the methylation pattern is lost; therefore, following bisulfite treatment, amplification of the gene of interest, and sequencing, no methylated cytosine should be detected within the sequence. These two controls can be used to monitor the effectiveness of immunoprecipitation technique.

To generate the artificially methylated genomic DNA, CpG methyltransferase (New England Biolabs, USA) was used according to the protocol provided. Briefly, 10xNEBuffer 2 (2 µl), S-adenosylmethionine (160 µM), genomic DNA (1 µg) and methyltransferase, *SssI* (40 units/µl; New England Biolabs, USA), were mixed in a final volume of 20 µl. The sample was incubated at 37°C for 4 hours followed by incubation at 65°C for 20 minutes. The final incubation step resulted in inactivation of the enzyme. This resulted in the generation of fully methylated genomic DNA.

To generate the un-methylated genomic DNA samples, DNA samples were sonicated for generation of 100-1000 bp fragments (SONICS Vibra Cell, 100 watt, 3 x 10 seconds with 35 seconds intervals on ice with 20% amplitude). The sonicated DNA samples were amplified using GenomePlex Complete Whole Genome Amplification (WGA) Kit (Sigma-Aldrich, Poole, Dorset, UK). Briefly, 10x fragmentation buffer (1 µl) was added to sonicated DNA samples (10 ng in a final volume of 10 µl) and samples were heated at 95°C for 4 minutes in a thermocycler. Following incubation, samples were immediately placed on ice for 3 minutes and consolidated by brief centrifugation. To generate the library, 1x library preparation buffer (2 µl) and library stabilisation solution (1 µl) were added to the samples and samples were thoroughly vortexed. Samples were heated at 95°C for 2 minutes and then placed on ice for 3 minutes. Library preparation enzyme (1 µl) was added to the samples while they were on ice. Samples were thoroughly vortexed, briefly centrifuged and then run on a thermocycler (Minicycler™, MJ Research, UK) with the following conditions: 16°C for 20 minutes, 24°C for 20 minutes, 37°C for 20 minutes and 75°C for 5 minutes. Following generation of the library, samples were placed on ice and a mastermix containing 10x amplification mastermix (7.5 µl), nuclease free water (47.5 µl) and WGA DNA polymerase (5 µl) were added to the samples. Samples were vortexed, briefly centrifuged then amplified using a thermocycler with the following conditions: 3 minutes at 95°C for the initial denaturation step, 14 cycles of 15 seconds at 94°C and 5 minutes at 65°C.

Fully methylated and un-methylated DNA samples were purified using QIAquick spin columns (Qiagen Ltd, West Sussex, UK), as described in section 2.5.1.1. Purified methylated and un-methylated genomic DNA samples were stored at -20°C until further use in MeDIP zebrafish tiling microarray experiments, described in Chapter 3.

2.6.6. DNA gel electrophoresis

DNA gel electrophoresis was used for optimisation of primers and separation of DNA molecules based on their size. Depending on the expected size of the products, the percentage of the gel varied from 0.6_2.5% (lower percentage gels were used for separation of larger DNA fragments while higher percentage gels were used for separation of smaller fragments). Agarose powder (Bioline Ltd., UK), according to the required percentage of the gel, was added to 1XTBE buffer (89 mM Tris base, 89 mM boric acid and 2 mM EDTA, adjusted to pH 8.0) containing ethidium bromide (0.5 µg/ml), and then dissolved by heating. When the temperature of the gel reached around 45°C it was poured into the tank and left to set. The gel was submerged in ethidium bromide containing 1X TBE Buffer (0.5 µg/ml). DNA samples and appropriate DNA molecular weight markers (100 bp, 1 kb; New England Biolabs, USA) were mixed with loading dye (6X gel loading dye, blue; New England Biolabs, USA) and then loaded into the wells. Electrophoresis was performed at 80 V for 20_45 minutes and the gel was visualised by UV transillumination.

2.6.7. Purification of DNA samples and sequencing

Prior to sequencing, PCR products were purified using QIAquick spin columns (Qiagen Ltd, West Sussex, UK), as described in section 2.5.1.1. The purified DNA samples were eluted in water, and then sequenced by the Functional Genomics and Proteomics Unit, School of Biosciences, University of Birmingham, Birmingham, UK using an ABI3730 DNA analyser (Applied Biosystems, UK). Each sequencing reaction was setup using 10 ng of the purified DNA samples and 0.4 µl of either reverse or forward primers (10 pmol). The final volume was adjusted to 10 µl using sterile water. To analyse the BSP data, the peak areas for C and T bases at each CpG site were measured and the methylation level was calculated using the formula below:

% methylated cytosine at each CpG site: $((C_{\text{peak area}})/(C_{\text{peak area}} + T_{\text{peak area}})) * 100$

2.7. Confirmation of the results of the dab oligonucleotide microarray using real-time PCR

To validate the microarray data, real time PCR (RT-PCR) was conducted on selected probes that their expression levels significantly varied between HCA and ST or HCA and healthy dab

liver samples according to microarray data. The same RNA samples that were used in the microarray experiments were also used for the RT-PCR experiment.

2.7.1. Primer design, validation and product sequencing

Primers were designed using Primer3 (Rozen and Skaletsky, 2000) and synthesised by AltaBioscience (Birmingham, UK). Primer sequences are presented in Table 2.4.

To synthesise cDNA, random hexamers (1 µl; Promega, UK) were added to 500 ng of DNase treated RNA samples (for DNase treatment method see section 2.5.3.5). Sample volumes were adjusted to 15 µl with RNase free water followed by incubation at 65°C for 2 minutes. The mixture was placed on ice prior to addition of a mastermix containing: 0.1M DTT (2 µl), 5x first strand buffer (5 µl), Superscript II reverse transcriptase (0.5 µl; Invitrogen, Paisley, UK), 25 mM dNTPs (0.5 µl; Fermentas Life Sciences, UK) and RNase free water (2.1 µl). The samples were run on a thermocycler with the following programme: 25°C for 10 minutes, 42°C for 90 minutes and 70°C for 15 minutes. The synthesised cDNA was quantified by NanoDrop ND-1000 UV-VIS Spectrometer version 3.2.1.

Primers were validated using DreamTaq DNA polymerase (Fermentas Life Sciences, UK). A mastermix containing: 10x DreamTaq buffer (5 µl), 5 µl dNTP mix (each dNTP at 0.2 mM), forward primer (10 pmol) and reverse primer (10 pmol), cDNA (80 ng), DreamTaq DNA polymerase (1.25 units) was prepared. The total volume was adjusted to 50 µl with water. The samples were run on a thermocycler with the following programme: 95°C for 3 minutes, 35 cycles of 95°C for 30 seconds, variable annealing temperatures for 30 seconds and 72°C for 30 seconds. Following completion of the cycling phase, a final extension step was performed at 72°C for 5 minutes.

PCR products were examined by agarose gel electrophoresis and products were confirmed by sequencing, as described in sections 2.6.6 and 2.6.7.

Gene name	Forward (5'→3')	Reverse (5'→3')	Product size (bp)
<i>vitellogenin b</i>	CTGACCTTCGTGGATATTGAG	ATCTGAGCCTCGGCATTG	145
<i>vitellogenin c</i>	AGTGAGGCAGACATGGTCCT	TCTCTGGCGGCTATGAGTCT	170
<i>S-adenosylhomocysteine</i>	TGATCTTCTCTGCTGCGTTC	TTCGACAACATGAAGGACGA	100
<i>udp-glucose 6-dehydrogenase</i>	AGACCTACGGGATGGGAAAG	CTTGGTGTTGGCATCGAATA	167
<i>syntrophin, beta 2</i>	5'-GGCCAGCTCCTGGTTCAC	AACCAGCCGATGTGTTTGAG	135

Table 2.4. Primer sequences for real-time PCR used for validation of the dab gene expression microarray results.

2.7.2. Real-time PCR

RT-PCR was conducted on an ABI Prism 7000 (Applied Biosystems, USA) using SYBR Green SensiMix (Quantace, Watford, UK). Five biological replicates, with three technical replicates of each were run. Each sample contained: cDNA (80 ng), Sensimix (12.5 µl), Sybergreen (0.5 µl), forward and reverse primers (depending on the primer efficiency values the amount of primers used ranged from 1 to 5 pmol) and nuclease free water (to adjust the final volume to 25 µl). The cycling parameters were: 95°C for 30 seconds (denaturing step) and 60°C for 30 seconds (combined annealing and extension). Melting curves were plotted using the ABI Prism 7000 SDS software to ensure only a single product was amplified and no primer dimers were formed. Absolute fluorescence values were used to calculate PCR efficiencies for each well (Ramakers et al., 2003). Threshold cycle (CT) values were recorded in the linear phase of amplification and the data were analysed using the delta-delta CT method of relative quantification (Livak and Schmittgen, 2001). The internal reference gene (*syntrophin*, *beta 2*) was selected as unvarying during the microarray experiment and was used to further normalise the data.

2.8. Targeted quantification of 12 metabolites related to the one-carbon cycle via liquid chromatography (LC)-triple stage quadrupole (TSQ) tandem mass spectrometry

Samples prepared for the targeted metabolomic study, described in section 2.2.2.3.2, were resuspended in methanol: water mixture (4 µl: 4 µl, HPLC grade water and methanol, J.T. Baker, Netherlands) and vortexed for 30 seconds. To concentrate the samples and increase the sensitivity of the detection, five aliquots of the same re-suspended samples were combined. The prepared samples were dried using a centrifugal evaporator (Thermo Savant, Holbrook, NY) and stored at -80°C until analysed. Prior to analysis, the concentrated samples were re-suspended in 7.5 µl buffer B (10 mM ammonium formate pH 2.2, 75% acetonitrile: 25% water, de-gassed and filtered, HPLC grade water and mass spectrometry grade acetonitrile, J.T. Baker, Netherlands) and 2.5 µl buffer A (10 mM ammonium formate pH 4.2, de-gassed and filtered). The re-suspended samples were vortexed and centrifuged for 10 minutes at 14,000 rpm. Samples (8.7 µl) were transferred to a 96 well plate. *S*-adenosyl-*L*-methionine-*d*3 (*S*-methyl-*d*3) tetra (*p*-toluenesulfonate) salt (CDN Isotopes, Essex, UK) was dissolved in methanol: water mixture (1:1) with a final concentration of 3.125 mmol/ml. The compound

was further diluted with buffer B: buffer A mixture (75%: 25%) to a final concentration of 0.126 mmol/ml. To each prepared sample (8.7 µl), 1 µl of the diluted *S*-adenosyl-*L*-methionine-*d*3 (*S*-methyl-*d*3) tetra (*p*-toluenesulfonate) salt was added to a final concentration of 0.016 mmol/ml. This compound was used as an internal standard.

The prepared samples were analysed using Dionex UltiMate 3000 liquid chromatography system with micropump coupled to a triple stage quadrupole (TSQ) tandem mass spectrometer (Thermo Fisher Scientific, Hertfordshire, UK) equipped with Ion Max-S atmospheric pressure ionisation (API) spray source (Thermo Fisher Scientific, Hertfordshire, UK). For technical replication, samples (1µl injections; acetonitrile was used for injection loop) were run in duplicates. Water: acetonitrile mixture (1:1; 1µl injections; acetonitrile was used for injection loop) was run in-between each set of biological samples to minimise carryover. Furthermore, to correct for background noise, negative controls were run at the beginning and at the end of the sample runs. Separation was achieved using a reverse phase column with weak anion exchange properties (Acclaim Mixed-Mode WAX column, 250 x 0.3 mm internal diameter, 5 µm particle size, 120 Å pore size, Dionex, Idstein, Germany) with column oven temperature of 20°C (minimum and maximum temperature of 18°C and 22°C, respectively) and under a gradient running buffer, as indicated in Table 2.5 and Figure 2.12. Analytes were detected with data acquisition performed as MRM. The masses of the precursor ions and the formed product ions, used for detection of the 12 metabolites of interest and the internal standard, are mentioned in Table 2.6.

The retention profiles of the compounds of interest were investigated by combining known amounts of the 12 metabolites of interest and the internal standard, *S*-adenosyl-*L*-methionine-*d*3, and injecting the achieved mixture into the LC-TSQ instruments. A mixture of the 12 compounds comprised of *S*-(5'-adenosyl)-*L*-homocysteine, *D*-methionine, folic acid, adenosine, betaine, sarcosine, *DL*-homocysteine, *S*-(5'-adenosyl)-*L*-methionine *p*-toluenesulfonate salt, glycine, *N,N*-dimethylglycine, choline *p*-toluenesulfonate salt, stachydrine hydrochloride (Extrasynthese, Genay, France) was prepared in water with each compound at a final concentration of 1 mmol/ml. The mixture was further diluted with buffer B (75%) and buffer A (25%) at a final concentrations of 0.05 mmol/ml per compound. The prepared mixture was spiked with the internal standard, *S*-adenosyl-*L*-methionine-*d*3 at a final sample concentration of 0.016 mmol/ml. As well as in combination, each standard was

individually directly injected into TSQ for establishment of the fragmentation patterns (product m/z) and the exact masses of the precursor ions to aid the identification of the compounds, as detailed in Table 2.6. Xcalibur 1.2 software (Thermo Fisher Scientific, Hertfordshire, UK) was used for measurement of the peak areas and analysis of the achieved data.

Steps	Retention time (minutes)	Flow (μ l/minutes)	%B	%C	%A
1	0	8	100	0	0
2	0	8	100	0	0
3	6	8	100	0	0
4	12	6	5	0	95
5	14	6	5	0	95
6	17	6	0	100	0
7	18	6	0	100	0
8	18.01	6	100	0	0
9	24	8	100	0	0
10	28	8	100	0	0

Table 2.5. The multi-step gradient used for liquid chromatography. Buffer B: 10 mM ammonium formate pH 2.2, 75% acetonitrile: 25% water, Buffer A: 10 mM ammonium formate pH 4.2, Buffer C: acetonitrile: water mixture (1:1).

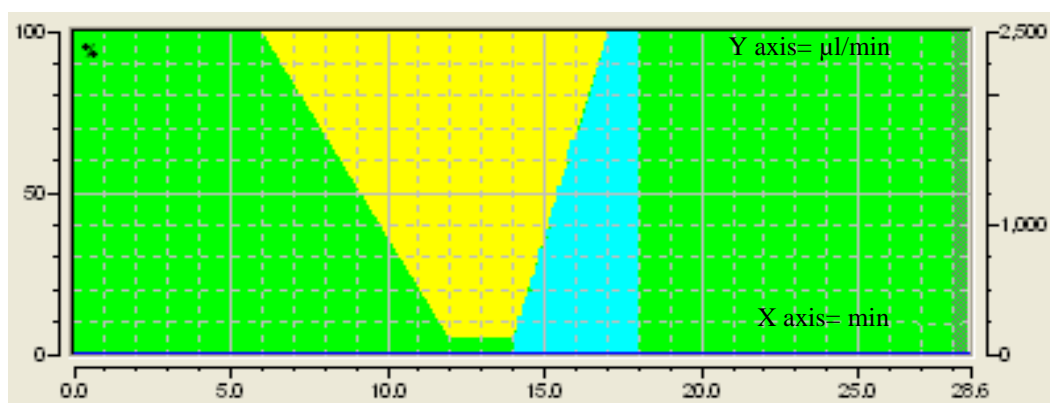


Figure 2.12. Schematic representation of the multi-step gradient used for liquid chromatography. Green: percentage of buffer B, Yellow: percentage of buffer A, Blue: Percentage of buffer C, X-axis: time (minutes), Y-axis: % Buffer.

Compound	Formula	Ion form	Parent mass	S-lens	Collision energy (%)	Product mass
<i>S</i> -(5'-Adenosyl)- <i>L</i> -homocysteine	C ₁₄ H ₂₀ N ₆ O ₅ S	[M+H] ⁺	384.937	106	37	87.962
					54	118.984
					19	133.95
					20	135.987
					13	232.036
					12	249.982
					12	249.982
					44	45.015
Homocysteine	C ₄ H ₉ NO ₂ S	[M+H] ⁺	135.952	55	32	47.033
					19	56.073
					18	73.013
					11	90.023
					5	118.012
					5	118.012
Stachydrine (Proline betaine)	C ₇ H ₁₃ NO ₂	[M+H] ⁺	143.973	93	36	42.056
					36	56.045

						25	58.052
						22	84.04
						20	98.047
						21	102.009
						21	102.009
						17	56.067
						23	61.022
						17	87.01
						12	102.022
						10	104.031
Methionine	C ₅ H ₁₁ NO ₂ S	[M+H] ⁺	149.953	58		6	133.006
						6	133.006
						57	105.964
						45	107.978
						34	119.947
						38	175.968
Folic acid	C ₁₉ H ₁₉ N ₇ O ₆	[M+H] ⁺	441.941	101		20	294.983

						20	294.983
						46	131.999
						37	175.022
						27	293.035
Folic acid	C ₁₉ H ₁₉ N ₇ O ₆	[M-H] ⁻	439.937	128		23	311.064
						24	396.195
						20	422.143
						20	422.143
						55	66.993
						54	91.976
						38	93.978
Adenosine	C ₁₀ H ₁₃ N ₅ O ₄	[M+H] ⁺	267.935	88		45	118.981
						19	135.993
						33	137.005
						33	137.005
Betaine	C ₅ H ₁₂ NO ₂	[M+H] ⁺	118	78		53	42.041
						45	43.049

						36	56.094
						26	58.045
						19	59.077
						63	28.092
						51	30.103
Sarcosine	C ₃ H ₇ NO ₂	[M+H] ⁺	89.985	43		25	42.064
						12	44.068
						25	59.037
						30	96.953
						24	101.917
<i>S</i> -(5'-Adenosyl)- <i>L</i> -methionine	C ₁₅ H ₂₃ N ₆ O ₅ S	[M+H] ⁺	398.944	110		28	136.022
						15	250.014
						15	263.892
						13	298.015
						62	28.087
Glycine	C ₂ H ₅ NO ₂	[M+H] ⁺	75.988	41		11	30.103
						45	42.009

Dimethylglycine	C ₄ H ₉ NO ₂	[M+H] ⁺	103.992	52	52	45.006
					19	47.016
					65	28.113
					29	30.114
					39	42.061
					39	44.072
					24	56.053
					14	58.06
Choline	C ₅ H ₁₅ NO ₂	[M+H] ⁺	104.028	70	42	30.109
					51	42.064
					21	45.06
					34	58.066
					22	59.07
					17	60.08
<i>S</i> -Adenosyl- <i>L</i> -methionine- <i>d</i> 3	C ₁₅ H ₂₀ D ₃ N ₆ O ₅ S	[M+H] ⁺	402.063	114	31	96.998
					25	102.044
					29	136.03

15	250.037
16	267.092
13	301.055

Table 2.6. The collision energy, S-lens value or stacked ring ion guide (SRIG) value (this value represents the radio frequency (RF) that efficiently captures and focuses the ions into a tight beam for transport of the ions from the source to the mass analyser), parent (precursor) masses, product masses, positive or negative ion mode and the chemical formula used for detection of the 13 compounds of interest are shown in this table.

2.9. Statistical analysis

2.9.1. General statistical approaches used throughout this thesis

SPSS v16 software was used for statistical analysis of the data. Normal distribution of the data and homogeneity of variance was evaluated via Sharipo-Wilks' test and Levenes' test, respectively. Data that were normally distributed with homogenised variance were analysed by one-way ANOVA with a Tukey post-hoc test (comparison of more than two groups with one factor) or by 2-tailed independent student's *t*-test. Data that did not meet the requirements for normal distribution or homogeneity of variance were analysed by non-parametric statistics, using a Kruskal-Wallis test (non-parametric test for comparison of more than two groups) and Mann-Whitney U test (non-parametric test for comparison of two groups). *P*-values <0.05, <0.01 and <0.001 are indicated by *, ** and ***, respectively.

2.9.2. Statistical analysis of microarray data

Analysis of zebrafish tiling microarray data:

GeneSpring v7.2 (Agilent) was used for analysing the data. Data flagged as present in at least 4 of the 10 samples were used for analyses. The Cy5 test signal was divided by the Cy3 control signal (genomic pool) and each microarray was normalised to the mean of the control group. Data with low raw intensity (less than 25) or standard deviation greater than 1.4 between biological replicates were removed from the analyses. Lists of hypo- and hyper-methylated genes were generated using greater than 1.5-fold change cut-off and parametric Welch *t*-tests with assumption of un-equal variance between tumour and healthy groups (*P* <0.05) on the Lowess normalised data. MIAME-compliant raw microarray data were submitted to ArrayExpress at EMBL-EBI and can be found under accession E-MTAB-209.

Analysis of cDNA flounder microarray data:

GeneSpring v7.2 (Agilent) was used for analysing the data. Data flagged as present in all the samples were used for analyses. The Cy5 test signal was divided by the Cy3 control signal (genomic pool) and each microarray was normalised to the mean of the control group. Data with low raw intensity (less than 20) or standard deviation greater than 1.4 between biological replicates were removed from the analyses. Lowess normalised data were imported into

MultiExperimental Viewer v4.7. Welch paired *t*-test with assumption of un-equal variance ($P < 0.05$) without multiple testing correction with a fold change cut-off of > 1.5 was used for identification of statistically significantly altered regions in HCA samples compared to non-cancerous surrounding tissues.

Analysis of the dab gene expression microarray data:

GeneSpring vGX11 (Agilent, UK) and MultiExperimental Viewer v4.7 were used for analysing the data. Agilent feature extraction software (FE) was used for quality control, both at sample and gene level. The samples that did not pass the 8 feature extraction values were removed from further analysis. In addition, below-background level entities were also removed from analysis. Data flagged as present or marginal in at least 3 out of 32 samples were used for analysis. Following background correction, log intensity and log ratio calculation, data were normalised by Lowess. Data were standardised by calculating the mean and standard deviation (SD) for each data point (x) in all samples (standardised value = $(x - \text{mean})/\text{SD}$). Statistically differentially expressed genes were identified using Welch paired *t*-tests between matching HCA and ST, one way ANOVA between the three groups (HCA, ST and healthy liver), and un-paired *t*-test between HCA and healthy liver ($P < 0.05$) with a fold change of greater than 1.5 and Benjamini-Hochberg (B-H) correction (Benjamini and Hochberg, 1995). MIAME-compliant raw microarray data were submitted to ArrayExpress at EMBL-EBI and can be found under accession A-MEXP-2084.

2.9.3. Ingenuity Pathway Analysis of the microarray data

The “Core Analysis” function of Ingenuity Pathway Analysis (IPA) (<http://www.ingenuity.com/>) was used to aid interpretation of the hypo- and hyper-methylated genes as well as under- and over-expressed genes (>1.5 -fold change). A B-H multiple testing correction (FDR $< 5\%$) was used to determine significant enrichment of annotations with biological functions and canonical pathways amongst these gene sets.

2.9.4. Principal component analysis and hierarchical clustering

The principal components analysis (PCA) feature embedded in MultiExperimental Viewer v4.7 and GeneSpring software were used in order to identify and summarise patterns within

gene expression and DNA methylation data. Hierarchical clustering was applied to dab transcription data using MultiExperimental Viewer v4.7 to further investigate distinct gene expression profiles between dab hepatocellular adenoma samples and ST samples.

Chapter 3

Comprehensive profiling of zebrafish hepatic proximal promoter CpG island methylation and its modification during chemical carcinogenesis

The data presented in this chapter has been published.
Mirbahai et al., (2011) *BMC Genomics*, 12 (3): 1-16.

3.1. Introduction

DNA methylation at CpG dinucleotides has a significant role in the regulation of gene expression (Rauch et al., 2007), as discussed in Chapter 1. This is achieved through modulation of protein-DNA interactions (Jones and Takai, 2001; Pomraning et al., 2009). Therefore, aberrant methylation of CGIs in the promoter and exon regions (Weber et al., 2005; Takai and Jones, 2002), and subsequent changes to gene expression can give rise to a range of disorders (Weber et al., 2005; Egger et al., 2004; Rauch et al., 2006).

Cancer is a complex and multifactorial disorder with more than 100 distinct types and subtypes found within specific organs in humans (Pogribny, 2010; Hanahan and Weinberg, 2000). Clarifying the molecular mechanisms of tumourigenesis and identifying fundamental common mechanisms between different types of tumours has been the aim of many researchers. Accumulation of both mutations and epigenetic abnormalities in the regulatory regions of genes and disruption of cellular signalling pathways contribute to tumour development (Tischoff and Tannapfel, 2008; Gronbaek et al., 2007). However, two key findings in regards to epigenetics and tumourigenesis have led to substantial interest in this field: 1. Reversibility of the epigenetic changes at early stages of tumourigenesis 2. Identification of a unifying theme in the development of cancers (the latter and the link between environment and epigenetic is explained in Chapter 4). Therefore, epigenomic profiling of cancer cells and investigating the involvement of epigenetic mechanisms in the development of tumours are amongst the most intensively studied fields in epigenetics in relation to human health (Vandegheuchte and Janssen, 2011). However, in contrast to the wealth of information available on epigenetic changes and development of cancer in humans; this area is substantially understudied in aquatic biology, especially in relation to disease.

Flatfish dab, used in biomonitoring within the UK waters, have a high prevalence of liver tumours with the occurrence at some sites exceeding more than 20% of sampled individuals (NMMP, second report, 2004; Stentiford et al., 2009; Small et al., 2010; Southam et al., 2008; Lyons et al., 2006). The causative agents and the molecular mechanisms leading to these tumours are unclear (discussed in Chapter 1 and 4). As a result, the main aim of this project was to investigate the involvement of DNA methylation abnormalities at the gene level and whole genome level in these tumours and to provide a testable hypothesis regarding the mechanisms involved in dab tumourigenesis. However, as dab's genome is not sequenced,

any comprehensive study at the gene level was challenging. Therefore, the involvement of DNA methylation in fish tumours was first investigated using the model fish species, zebrafish (*Danio rerio*). This was the first of such studies on fish tumourigenesis.

Endogenous and exogenous factors (e.g. chemicals, viruses, radiation) are known to induce tumours (Loeb and Harris, 2008; Miller et al., 1970). A range of organic and inorganic chemicals in the environment have been shown to initiate and promote tumourigenesis in various species and tissues. Therefore, chemical induction of tumours is a useful tool for studying molecular mechanisms of tumourigenesis (Miller et al., 1970). In the last few decades, alternative models to rodents, such as fish, have been used extensively for studying the effects of environmental carcinogens. Several factors, including sensitivity to a variety of carcinogens, low cost of maintenance, short reproductive cycle, possibility of field studies, portability, potential for lifetime bioassays and the ease of genetic studies have made them ideal for initial carcinogenicity studies (Bailey et al., 1996; Spitsbergen et al., 2000). For example, rainbow trout (*Oncorhynchus mykiss*) (Bailey et al., 1996), zebrafish (*Danio rerio*) (Khudoley, 1984), guppy (*Poecilia reticulata*) (Hawkins et al., 1988), common platyfish (*Xiphophorus maculatus*) (Schwab and Scholl, 1981) and medaka (*Oryzias latipes*) (Hawkins et al., 1988) have been employed in bioassays for carcinogens. Zebrafish are a well-established model for investigating embryogenesis, organogenesis, and environmental carcinogenicity and for modelling human diseases such as cancer (Bailey et al., 1996; Wardle et al., 2006; Berghmans et al., 2005; Zon, 2010; Lam and Gong, 2006; Lam et al., 2006). Chemically-induced tumours are histopathologically very similar between zebrafish and humans (Berghmans et al., 2005; Amatruda et al., 2002) and orthologous TSGs and oncogenes in human and fish have been identified (Berghmans et al., 2005). Recent studies comparing hepatic gene expression in human and zebrafish tumours have demonstrated conservation of gene expression profiles at various stages of tumour aggressiveness between these two phylogenetically distant species (Lam et al., 2006). However, the contribution of altered DNA methylation to such changes was unknown.

Although there are reports on global levels of DNA methylation in zebrafish during embryogenesis or in adults (MacLeod et al., 1999; MacKay et al., 2007; Martin et al., 1999; Mhanni and McGowan, 2004; Jabbari and Bernardi, 2004), the effects of temperature on global levels of DNA methylation (Varriale and Bernardi, 2006a), DNA methylation profiling

in zebrafish embryo (Fenga et al., 2010), and the roles of chromatin-mediated gene regulation during embryogenesis (Wardle et al., 2006), no study had been published on the DNA methylation patterns or DNA methylation changes associated with carcinogenesis in zebrafish.

The zebrafish is already extensively employed in environmental carcinogenicity studies, as it can be easily raised and treated in a controlled laboratory environment and its genome sequence is readily available. Therefore, zebrafish was chosen as a subject for this investigation. Hepatocellular carcinoma (HCC) tumours, chemically induced by 7,12-dimethylbenz[α]anthracene (DMBA) in zebrafish, were used to investigate the differences in DNA methylation profiles between fish liver tumours and healthy fish livers. In addition, although zebrafish are commonly used for modelling human cancers, there is a lack of information on the suitability of zebrafish as an epigenetic model. Therefore, this study could also provide background information for future use in zebrafish models of human cancer.

In the past few years substantial progress has been made in the development of genome-wide DNA methylation techniques. In this chapter, a genome-wide approach, known as methylated-DNA immunoprecipitation (MeDIP)-microarray, was used for profiling DNA methylation alterations in zebrafish liver tumours. This method was initially used in a study conducted by Weber et al., (2005) for DNA methylation profiling of human cancer cells. MeDIP relies on un-biased enrichment of methylated fragments of DNA which can be combined with any type of DNA microarray. In this experiment a novel zebrafish tiling microarray was designed covering the regions of 1.5kb upstream to 1kb downstream of the transcription start site (TSS) of CpG island-containing genes. This was combined with MeDIP to achieve a comprehensive DNA methylation profiling. This study is the first to describe MeDIP in combination with a custom-made CGI tiling microarray for establishing normal and HCC DNA methylation profiles in the liver of any adult fish species. This microarray, in combination with well-established methylated-DNA immunoprecipitation, serves as a powerful tool for elucidating comprehensive DNA methylation profiles.

This work was carried out in collaboration with Professor Zhiyuan Gong's research group at the University of Singapore. The chemically induced zebrafish liver tumour samples were provided by our collaborators.

3.2. Overview of the experimental approach (Methodological details are described in Chapter 2)

Three different methods; HPLC, LC-MS/MS and an ELISA colourimetric technique, were investigated to identify a suitable methodology for measurement of the overall DNA methylation levels in fish liver.

To investigate gene specific DNA methylation changes in fish tumours, chemically induced zebrafish HCC samples were used. To induce HCC tumours, three-week-old zebrafish fry were treated with DMBA (0.75 ppm in fish tank water) dissolved in dimethyl sulfoxide (DMSO, vehicle) or DMSO alone (vehicle control) for 24 hours and the treatment (1.25 ppm DMBA or DMSO in fish tank water) was repeated once, 2 weeks later, for another 24 hours. Treated fish were rinsed three times in fresh water and transferred into new tanks for maintenance. Fish were sampled 6_10 months after DMBA exposure. HCC samples used for this study were all confirmed by histopathology studies.

To investigate differences in DNA methylation profiles in HCC samples compared to control samples, a custom-made CGI tiling microarray was developed. DNA samples from both groups were sonicated, immunoprecipitated with 5-methylcytosine antibody, labelled with different fluorescent dyes and hybridised to the microarray. GeneSpring, Blast2GO and Ingenuity Pathway Analysis were used to identify the biological functions that were altered in HCC samples compared to controls.

Bisulfite sequencing PCR was used to confirm the data obtained from the MeDIP-tiling microarray experiment.

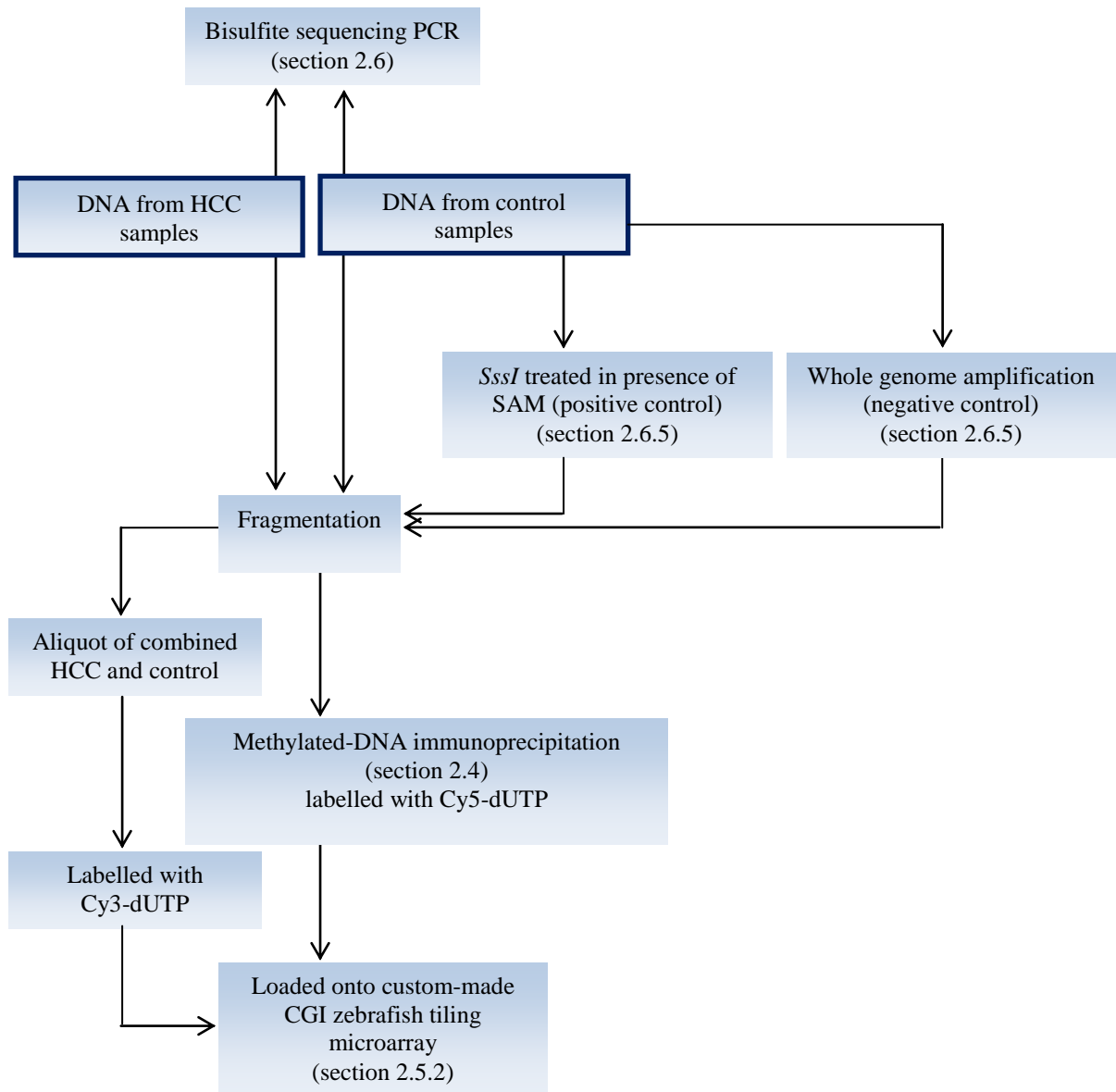


Figure 3.1. Flowchart of the experimental procedures used in Chapter 3. The starting points of the diagram are the boxes with borders. The methods used in this chapter are described in detail in Chapter 2. In this flowchart, the sections where these methods are described in Chapter 2 are shown in brackets.

3.3. Results

3.3.1. Global measurement of genomic DNA methylation

3.3.1.1. Establishing a method for measuring global genomic DNA methylation levels

Three different methods; HPLC, LC-MS/MS and ELISA, were compared for accuracy and suitability for measuring global genomic DNA methylation levels. In the ELISA method, (as described in section 2.3.1) DNA is immobilised to the surface of a plate. 5-methylcytosine antibody, detector antibody and a developing solution is added to the DNA. The amount of methylated DNA is proportional to the OD intensity. However, in the other two methods (as described in section 2.3.2) quantification of the methylated DNA is based on comparison to standards.

To test the three methods, the overall methylation levels in the livers of healthy dab and flounder were measured. As shown in Figure 3.2 no significant differences were detected between the three methods in both species investigated. However, as indicated by the error bars the variation in the colourimetric ELISA-based method was higher than the other two methods. This could be due to lack of antibody specificity, slight variation in efficiency of the washing steps, variation in the length of time taken to add the developing reagent and to measure the fluorescent signal between different samples and substrate availability. The HPLC technique used is an established in-house method. However, LC-MS/MS was carried out in an external laboratory as described in Chapter 2, section 2.3.2.7. The latter was mainly used as an external validation method and to compare the reproducibility and accuracy of the global DNA methylation measurements. As shown in Figure 3.2, HPLC and LC-MS/MS data were similar with low levels of variation between different samples. Therefore HPLC was used as the standard methodology for quantification of global DNA methylation levels throughout this project.

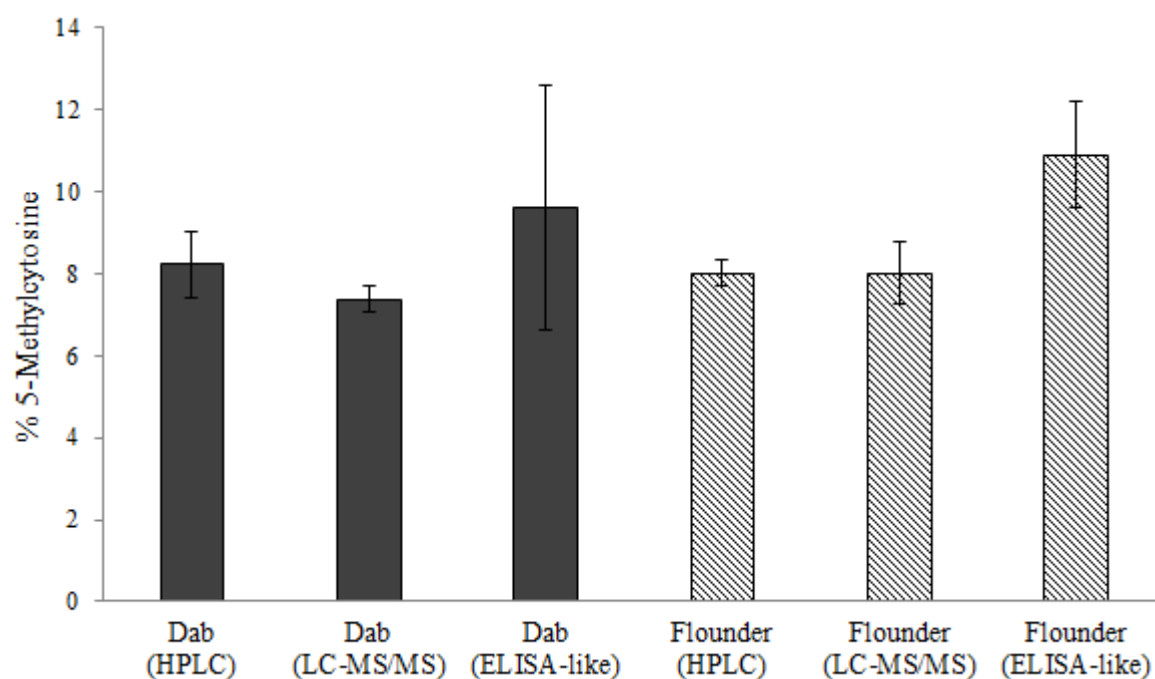


Figure 3.2. Comparison of HPLC, LC-MS/MS and ELISA methods for measurement of global levels of methylation in DNA. The same samples were used for all three techniques. Dab liver: n=3 different fish samples \pm standard error of the mean (SEM). Flounder liver: n=3 different fish samples \pm SEM. Data are normally distributed with homogenised variance. No significant differences were detected between the three techniques as determined by one-way ANOVA with Tukey's post-hoc test.

3.3.1.2. Comparison of global genomic DNA methylation in fish and mammals using HPLC

To establish the global levels of cytosine methylation in fish in comparison to mammals, HPLC was performed. To establish the retention times and the order of elution, mononucleotide standards –dCMP, 5mdCMP, dGMP, dTMP, dAMP- were run individually and as a mixture at least three times. The retention times and the order of appearance are shown in Figure 3.3.A.

RNA contamination was enzymatically removed from the samples as described (section 2.3.2.1). To confirm the absence of RNA contamination, a standard for uridine monophosphate (UMP) was used (Figure 3.3.B). DNA extracted from livers of zebrafish, dab, flounder and commercially available DNA from calf (*Bos taurus*) thymus were prepared as described (Section 2.3.2.2) and analysed using HPLC. As shown in Figure 3.4, the overall DNA methylation levels did not differ between the three fish species investigated with 5-methylcytosine comprising 8% of the total cytosine in the DNA. However, all three fish species had statistically significant ~2-fold higher DNA methylation levels ($P<0.05$) than as found in calf thymus DNA. This is in agreement with previous published data showing higher levels of DNA methylation in fish compared to mammals (Jabbari and Bernardi, 2004; Aniagu et al., 2008).

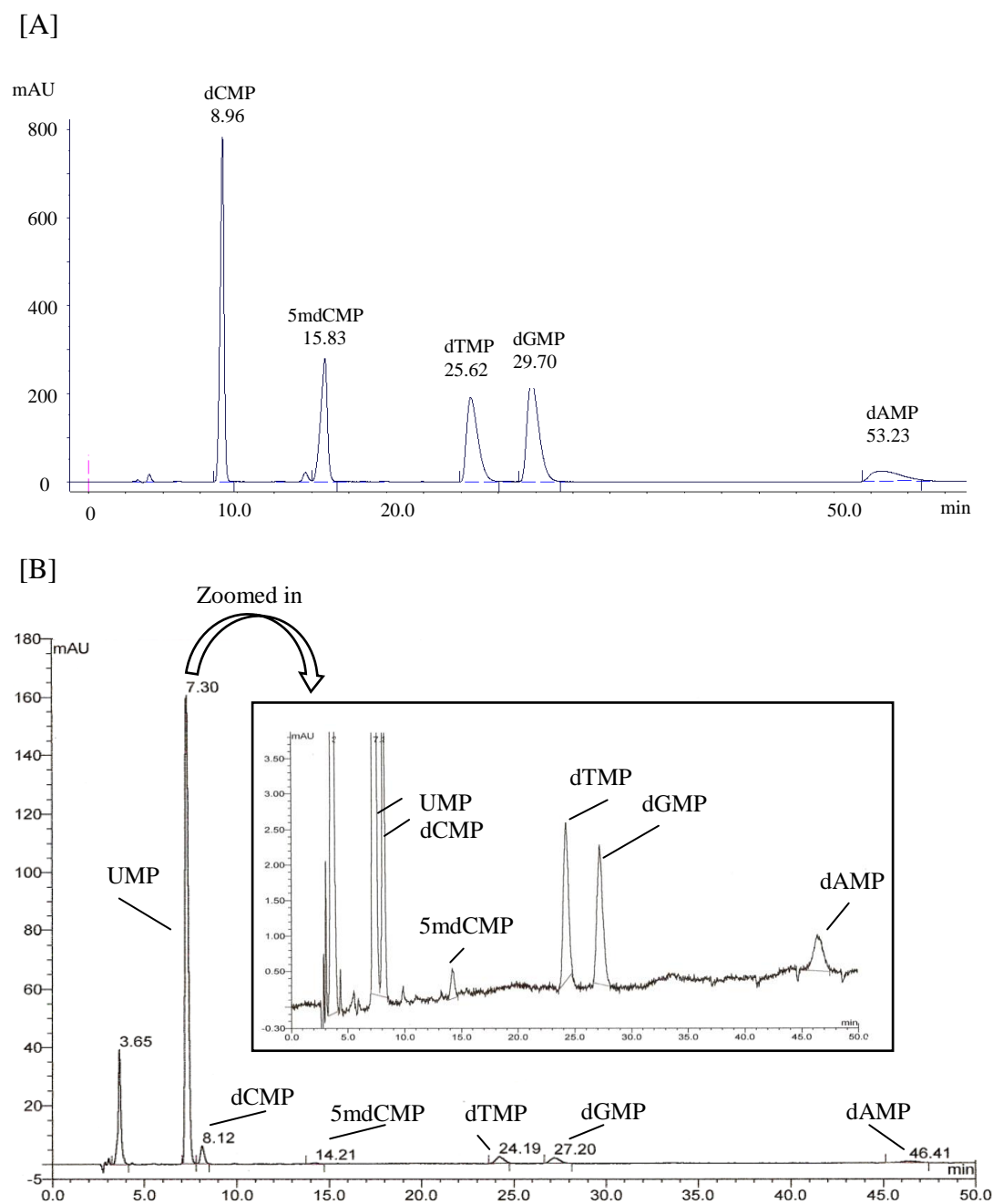


Figure 3.3.A. Determination of the retention times of five standard mononucleotides using HPLC. B. Spiked standard mixture with uridine monophosphate (UMP) for establishing the retention time of UMP for monitoring RNA contamination.

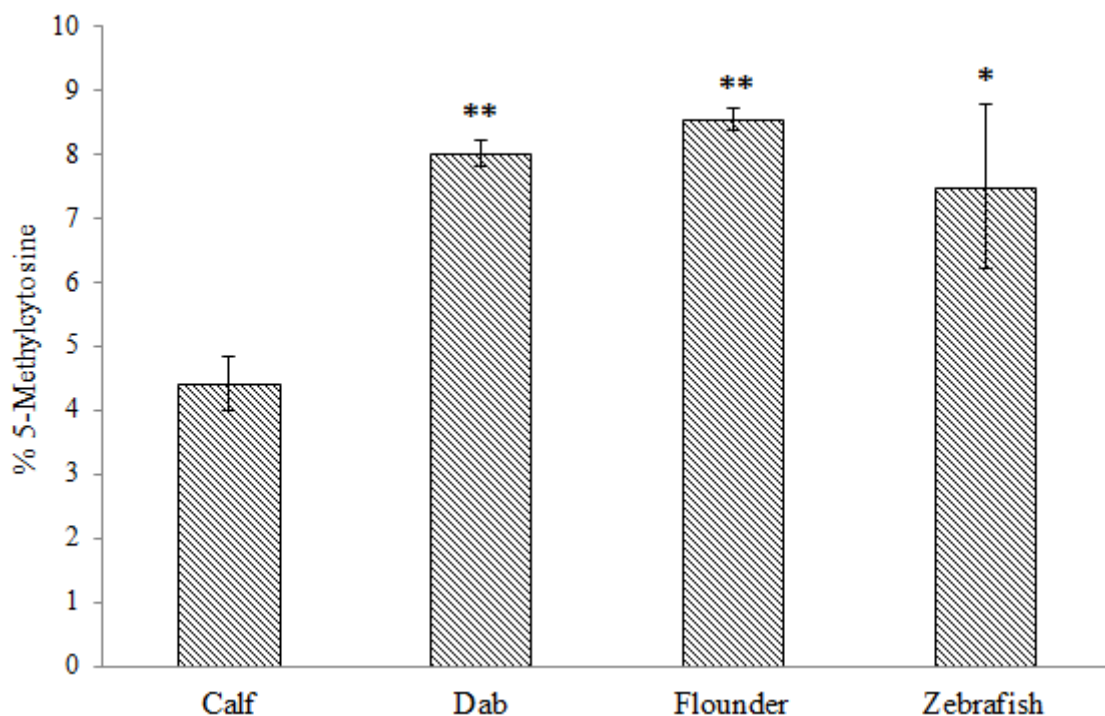


Figure 3.4. Measurement of the global percentage of methylated cytosine in calf thymus, adult male zebrafish, immature flounder and adult female dab livers using HPLC. No differences were detected between fish species investigated. The global percentage of methylation was statistically significantly higher (~2 fold) in all fish species compared to calf. Data are the mean \pm SEM of four independent experiments. ** and * statistically significantly different from calf thymus with $P < 0.01$ and < 0.05 , respectively. P -values were determined using a one-way ANOVA with Tukey's post-hoc test.

3.3.2. DNA methylation at gene levels in zebrafish liver tumours

In this section, differences in methylation levels of genes in DMBA-induced adult zebrafish HCC samples were established compared to control zebrafish liver samples. This was achieved by combining MeDIP with the custom-made zebrafish DNA tiling microarray.

3.3.2.1. Unbiased enrichment of methylated DNA using methylated-DNA immunoprecipitation (MeDIP)

The immunoprecipitation method as described (Section 2.4) was used for enrichment of methylated fragments of DNA extracted from both normal adult zebrafish liver and chemically induced HCC samples. DNA samples were extracted from tissue samples and

kindly provided by Professor Gong's research group as described in Chapter 2, section 2.2.1 and in their publication (Lam et al., 2006). To test the specificity of the immunoprecipitation method, CGI regions of 14 genes (15 regions) located 1.5kb upstream to 1kb downstream of transcriptional start sites were screened for methylcytosine using bisulfite sequencing PCR. These genes included: Transcriptional regulator *myc-b* (*mycb*, 2 regions), *cadherin 1* (*cdh1*), *cell division protein kinase 8* (*cdk8*), *timp metalloproteinase inhibitor 2* (*timp2*), *heat shock protein 90-beta* (*hsp90b1*), *mitogen-activated protein kinase 1* (*mapk1*), *tumour suppressor p53* (*p53*), *muts homolog 6* (*msh6*), *muts homolog 2* (*msh2*), *DNA (cytosine-5-)-methyltransferase 4* (*dnmt4*), *rap2c*, *member of ras oncogene family* (*rap2c*), *connexin-43* (*cx43*), *glutathione S-transferase P1* (*gstp1*) and *no-tail a* (*ntla*).

Following bisulfite treatment and sequencing of the 14 genes, CGIs located at the promoter region of the *ntla* gene were found to be methylated while CGIs at the remaining 13 genes, including *gstp1*, were completely un-methylated (Figures 3.5 and 3.6). The *ntla* gene and the *gstp1* gene were thus used as methylated (positive control) and un-methylated (negative control) DNA regions for testing the immunoprecipitation enrichment method. This was in addition to artificially methylated and un-methylated genomic DNA used as controls in the experiment.

DNA from one zebrafish liver sample was sonicated to generate 200 to 1000 bp enriched fragments. Fragmented DNA samples were immunoprecipitated using 5-methylcytosine (5-mC) antibody and bisulfite-treated. The immunoprecipitation resulted in enrichment of methylated fragments of DNA while the bisulfite treatment resulted in distinction of methylated cytosines from un-methylated cytosines in the enriched fragments. Equal amounts of immunoprecipitated, bisulfite-treated DNA were amplified for the two genes of interest. The gel images provided a semi-quantitative indication of enrichment of methylated *ntla* gene in comparison to un-methylated *gstp1* gene using 5-methylcytosine antibody (Figure 3.7). This was further confirmed by sequencing.

A:

ATCTCGTGCA CGCGTGTGGTATGAATGCAACGACAAACACAATTGCGTGCTCGCGGGGCGCGTGATGA

B:

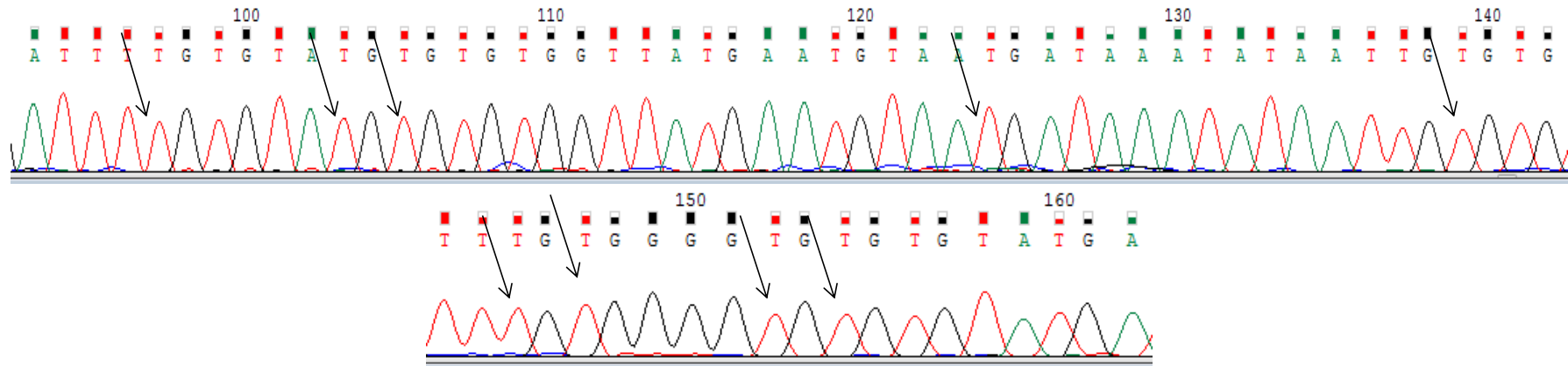
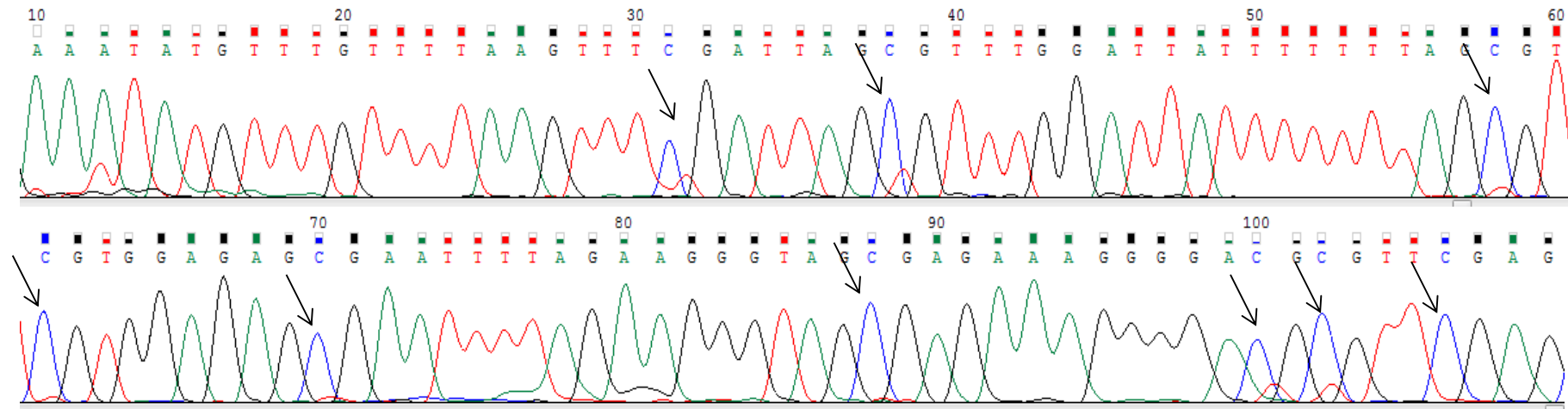


Figure 3.5. Direct bisulfite sequencing of *glutathione S-transferase P1* gene (*gstp1*). CpG sites investigated were located in a CGI located -98 to +19 from the transcriptional start site of *gstp1* gene. A. The original sequence (prior to bisulfite treatment) for the region investigated for *gstp1* with bisulfite sequencing PCR. CpG dinucleotides are shown in red and cytosines located in regions other than CpG dinucleotides are shown in green. The latter is used as internal controls for monitoring bisulfite treatment efficiency. B. As shown in this figure all CpG sites were un-methylated as all CpG dinucleotides were converted to TpG. Sequencing was performed on 4 independent samples; two tumour samples and two control samples. Arrows: thymines located at CpG dinucleotide sites that were converted by bisulfite treatment from un-methylated cytosines.

A:

AAATATGTCTGCCTCAAGTCCCGACCGAGCGCCTGGATCATCTCCTTAGCGCCGTGGAGAGCGAATTTCTAGAAGGGC
AGCGAGAAAGGGGA~~CGCGT~~CGAG

B:



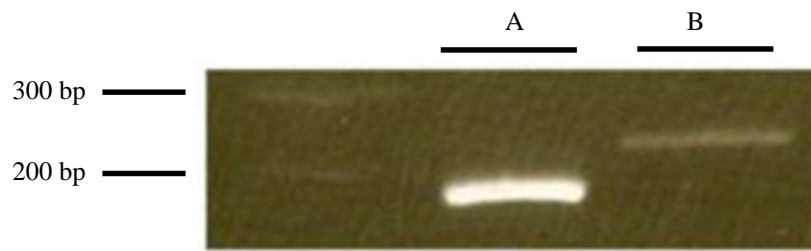


Figure 3.7. Agarose gel electrophoresis image of methylated-DNA immunoprecipitated (MeDIP) bisulfite treated *ntla* and *gstp1* genes. The methylated *ntla* (A) gene is enriched after immunoprecipitation compared to un-methylated *gstp1* (B) gene.

3.3.2.2. Design of the zebrafish CGI tiling microarray and comprehensive mapping of adult zebrafish CGI methylation

Genes with CGIs located 1.5 kb upstream to 1kb downstream of the TSS were identified as described (section 2.5.2.1). Probes (43,960) were designed to cover the identified 6,024 promoters of the genes containing CGIs. A list of genes with predicted CGIs, numbers of CGIs for each chromosome and location on the chromosome are presented in Appendix, File 1. Where possible, overlapping 60–mer probes, with 25 bp spacing, were designed to cover the entire length of CGIs (Figure 3.8). The microarray design was named Agilent Birmingham *D. rerio* 025794 45220v1 and is available from ArrayExpress under accession A-MEXP-1813. The tiling microarray provided a detailed, un-biased approach for studying DNA methylation (Mockler and Ecker, 2005).

Negative and positive controls were used to monitor the success of DNA methylation immunoprecipitation. The negative control consisted of PCR amplified whole genomic DNA, fragmented, immunoprecipitated and labelled with Cy5. As DNA methylation is removed during amplification, as expected, little DNA was precipitated during MeDIP. Therefore, a uniform low intensity signal was measured for all probes represented on the tiling microarray. This represented non-specific binding of DNA to the 5-methylcytosine antibody. The positive control was achieved by treatment of DNA with CpG methyltransferase, *Sss I*, in the presence of *S*-adenosylmethionine (SAM) prior to MeDIP. This resulted in complete methylation of all CGIs and detection of a uniform, high fluorescent signal for all probes (Appendix File 2,

Table 1). The positive control, in addition to serving as a control for the microarray experiment, was used to monitor the efficiency of MeDIP (Figure 3.9)

The CGI tiling microarray provided a comprehensive mapping of the DNA methylation at CGIs in healthy adult zebrafish liver between 1.5 kb upstream and 1kb downstream of transcription start site of the genes (Figure 3.10). The list of all genes on the tiling microarray and their methylation levels are presented in the Appendix (File 2, Table 4. those genes with values 2-fold above the median level (higher DNA methylation levels) and 2-fold below the median level (lower DNA methylation levels) in control samples are highlighted). Gene ontology analysis using Blast2GO identified GO terms significantly (false discovery rate (FDR) <5%) over-represented amongst genes that were relatively low and also relatively high in DNA methylation levels in healthy zebrafish liver. For genes with lower levels of methylation in healthy zebrafish liver, molecular functions associated with biological processes related particularly to involvement in regulation of gene expression and transcription, cell cycle, negative regulation of apoptosis, intracellular protein transport and mitochondrial transport, developmental process, mitochondrial differentiation, cellular response to hormone stimulation and sex determination were enriched (due to the size of the image and clarity, the complete list of biological processes enriched in lower methylated regions are shown in Appendix, File 3 with some of the enriched biological processes shown in Figure 3.11). For genes with higher levels of methylation in healthy zebrafish liver, molecular functions, such as molecular transducer activity, signalling transducer activity and receptor activity associated with the G-protein coupled receptor protein signalling pathway were enriched (Figure 3.12).

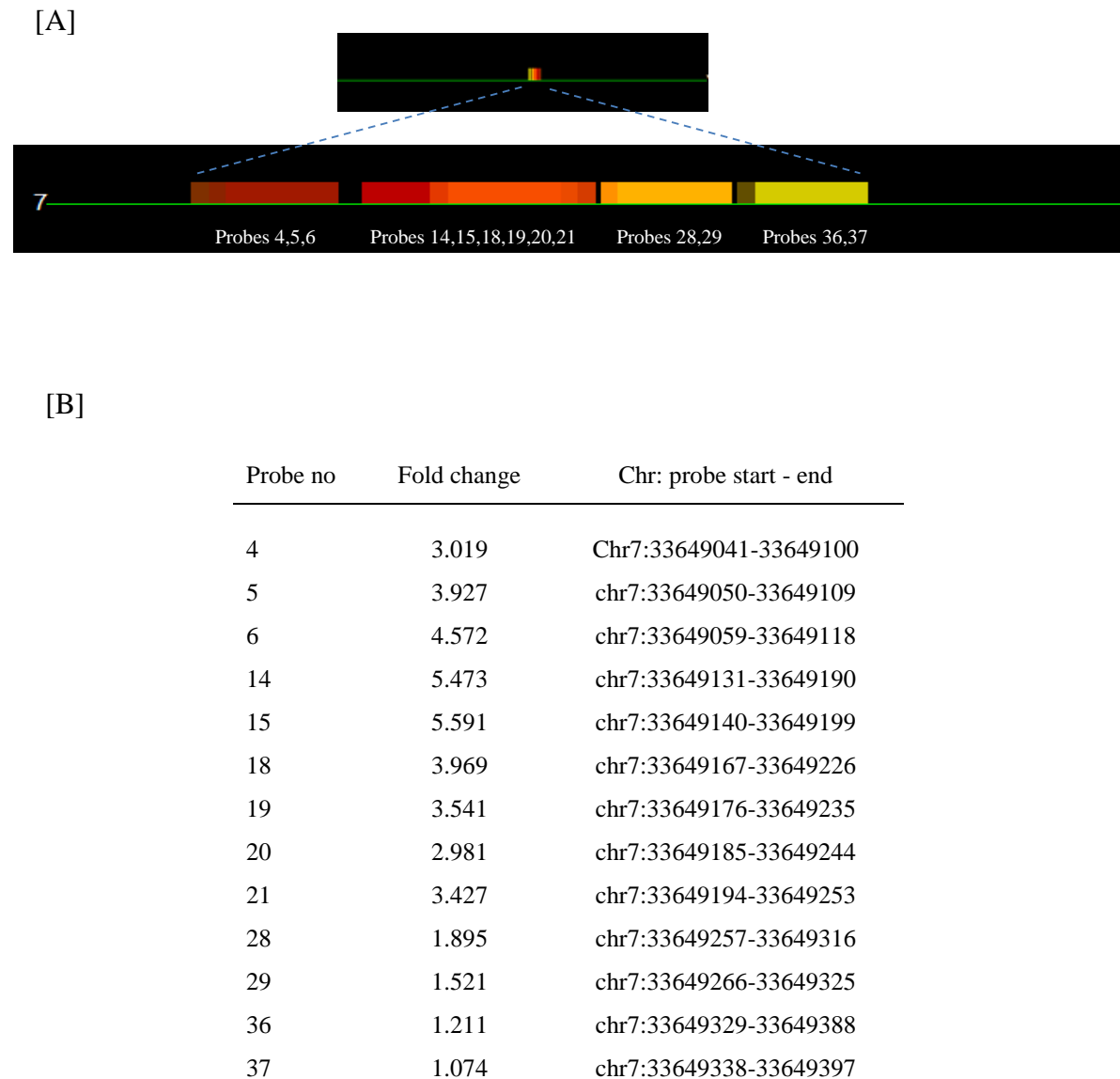


Figure 3.8. Chromosomal mapping of the probes. A. All probes were mapped onto zebrafish chromosomes. As an example, probes for *coronin*, *actin binding protein 2ba* gene (*coro2ba*; ENSDARG00000079440) on chromosome 7 are shown. A CGI was predicted at the region between 33649014 and 33649398. Colours indicate the intensity of the fluorescent signal from the probes (red: hypermethylation, yellow: no change, blue: hypomethylation). B. The mean fold changes for all probes representing the *coro2ba* gene in tumour samples compared to healthy zebrafish liver samples.

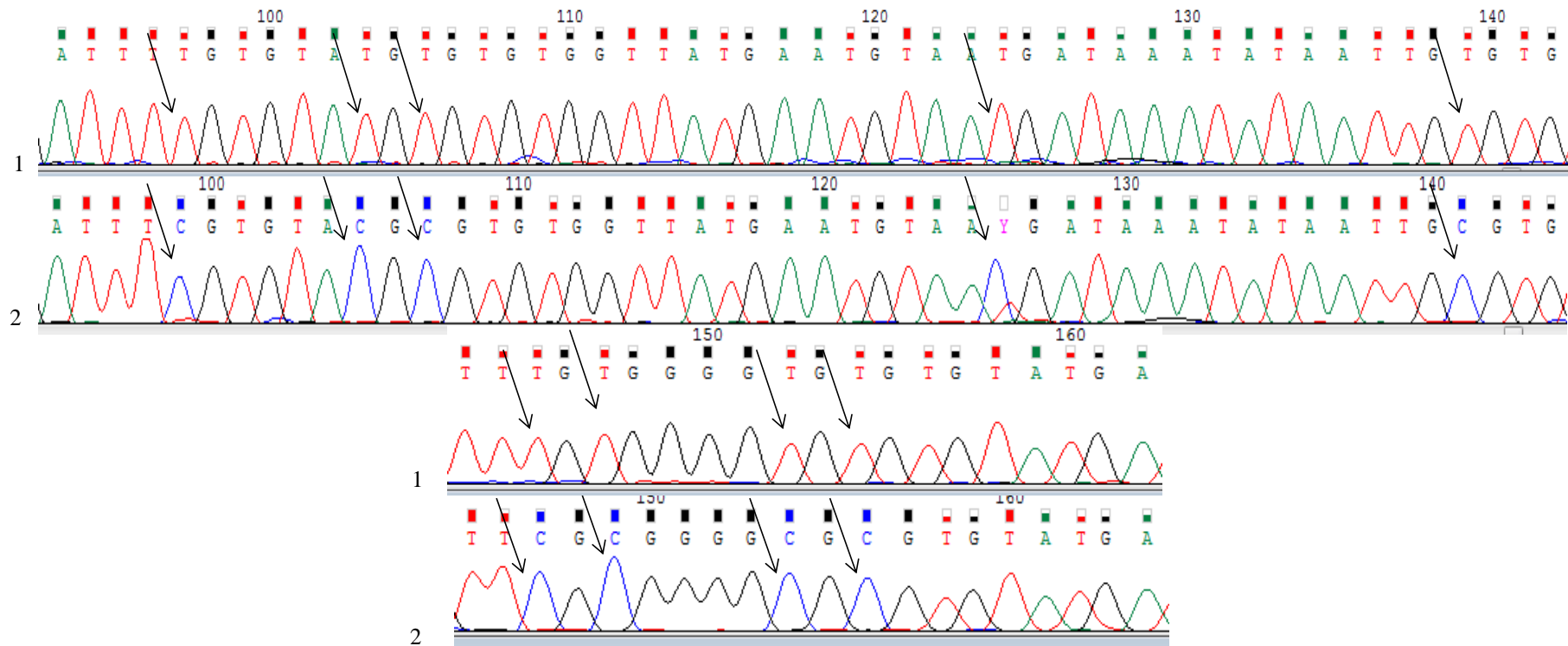


Figure 3.9. Generation of artificially methylated genomic DNA. *Glutathione S-transferase P1*, as explained in Figure 3.5, is un-methylated in both tumours and healthy zebrafish livers. Treatment with CpG methyltransferase, *SssI*, in the presence of *S*-adenosylmethionine (SAM) results in methylation of all CpG dinucleotides. Successful bisulfite treatment of the positive control was established by sequencing and observing a single C peak at all CpG dinucleotide positions. 1: *gstp1*, 2: *SssI* treated *gstp1*. Arrows: CpG dinucleotides.

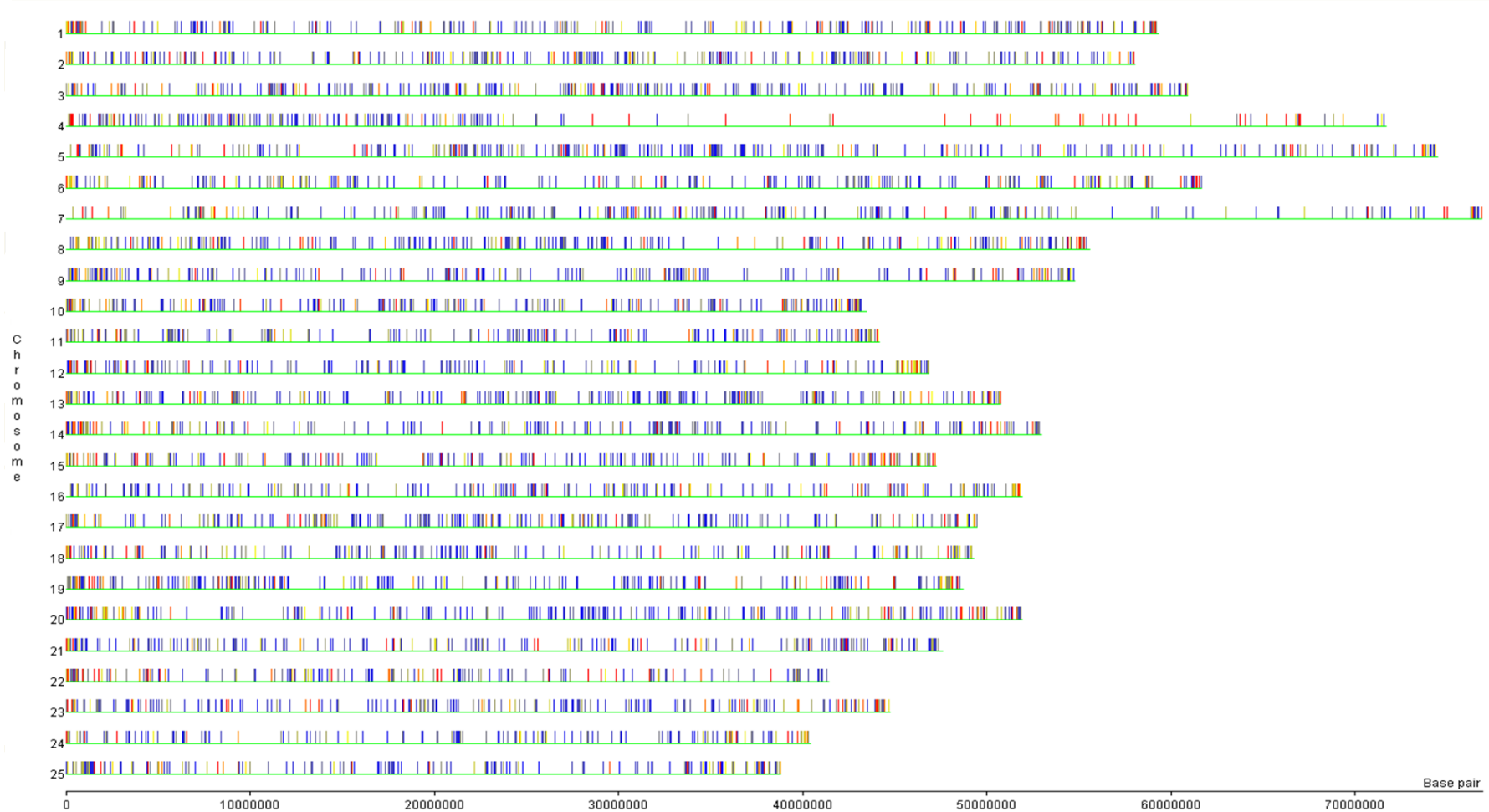


Figure 3.10. Chromosomal mapping of average methylation levels in tumour samples compared to control samples. Chromosomes are numbered 1_25. Each line on the chromosome indicates a probe on the zebrafish microarray. Red: hypermethylation, Blue: hypomethylation, yellow: no change.

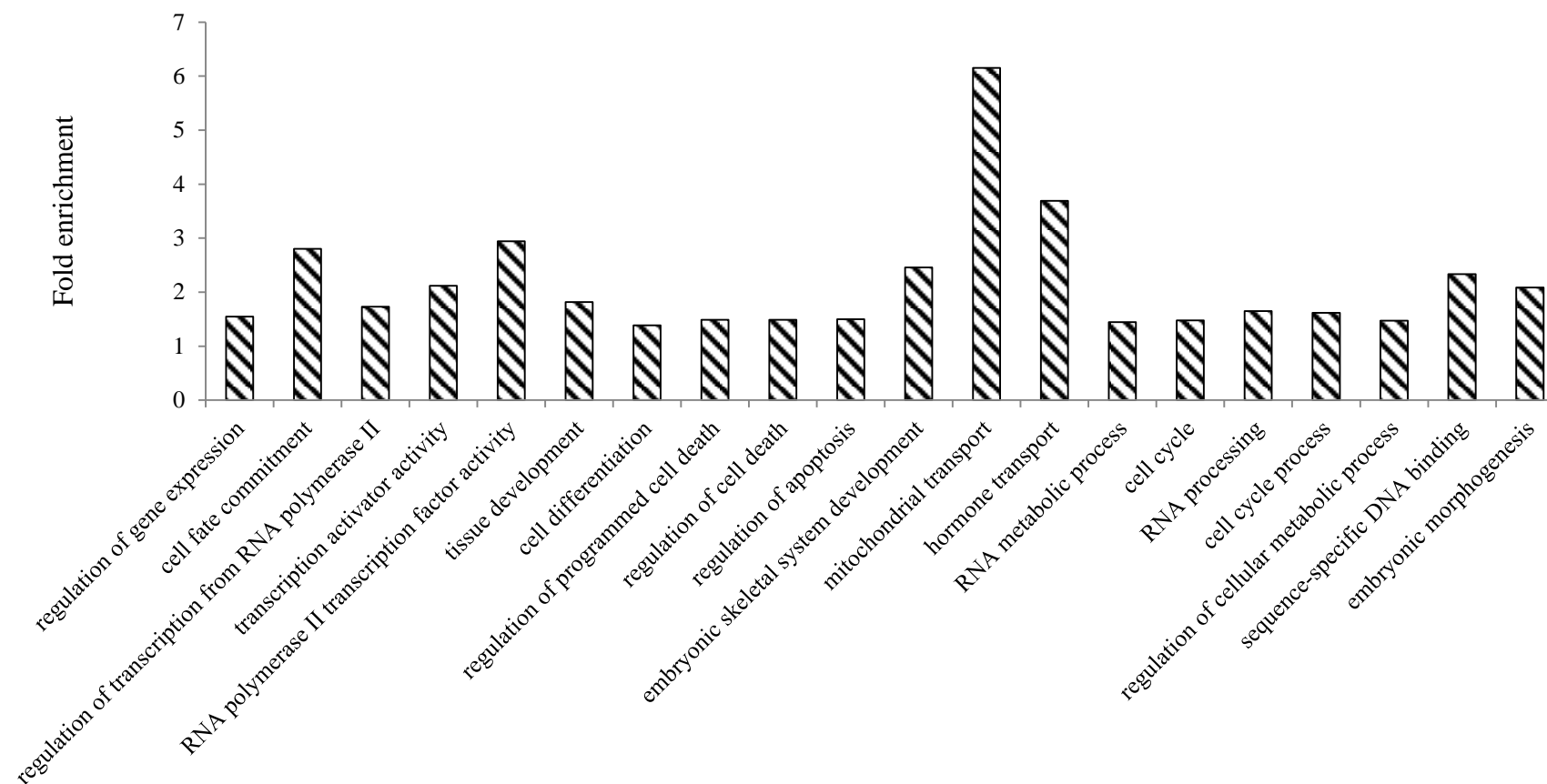


Figure 3.11. Selected gene ontology (GO) terms significantly over-represented in the genes with DNA methylation levels 2-fold below the median level in healthy zebrafish liver i.e. genes with lower methylation levels (FDR < 5%). The complete list of biological processes enriched in lower-methylated regions is shown in Appendix, File 3.

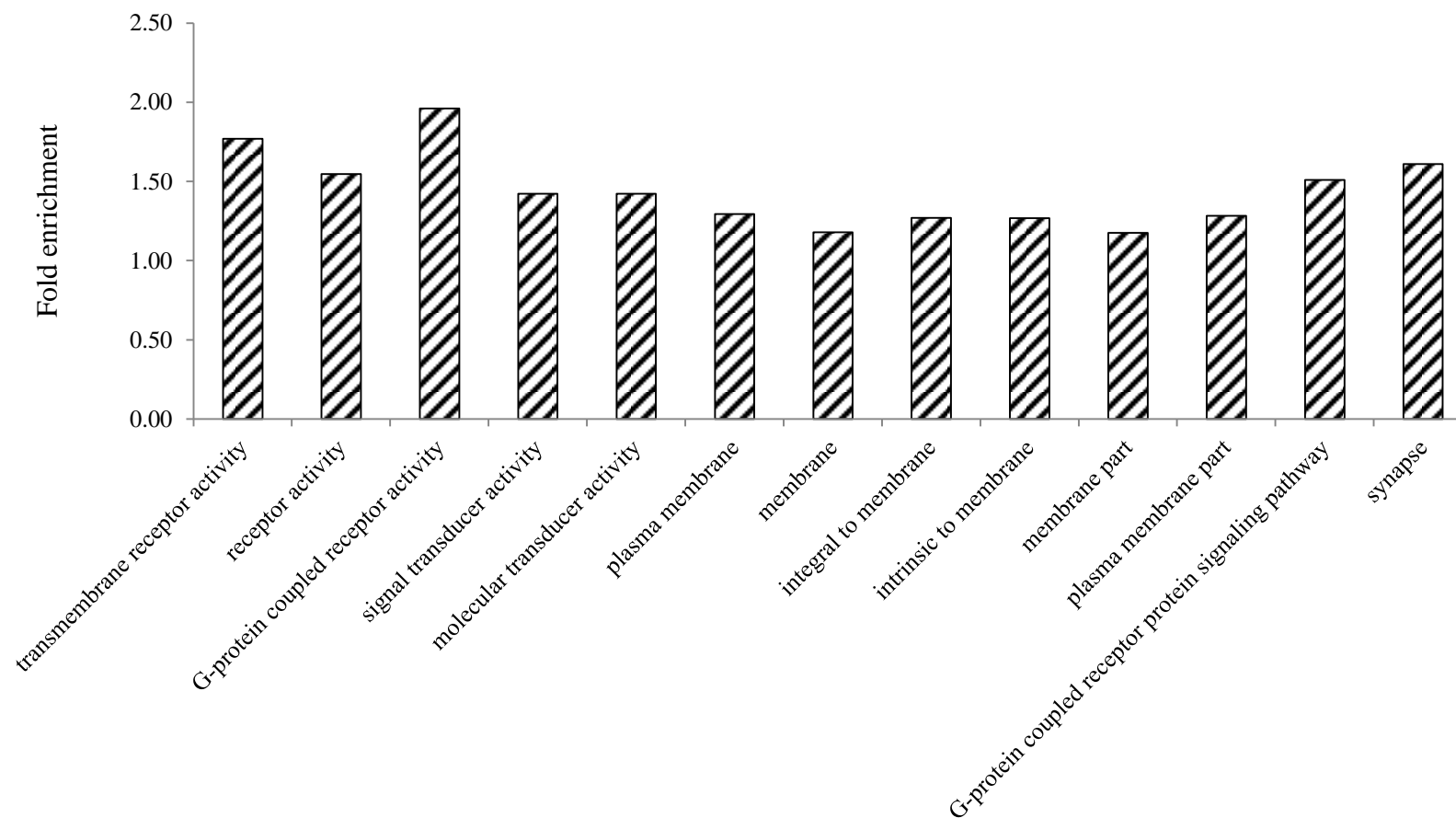


Figure 3.12. Gene ontology (GO) terms significantly over-represented in the genes with DNA methylation levels 2-fold above the median level in healthy zebrafish liver i.e. genes with higher methylation levels (FDR< 5%).

3.3.2.3. DNA methylation analysis of zebrafish hepatocellular carcinoma and comparison to gene expression

Genomic DNA extracted from zebrafish HCC samples was processed by MeDIP, labelled with fluorescent Cy5-dUTP and hybridised to the tiling microarray against input genomic DNA that had been labelled with Cy3-dUTP. Following quality checks and normalisation (as discussed in Chapter 2, section 2.9.2), probes were mapped onto the chromosomes and two lists, containing hypomethylated and hypermethylated regions in comparison to control, were generated (>1.5 -fold change, $P < 0.05$). MIAME-compliant raw microarray data were submitted to ArrayExpress at EMBL-EBI and can be found under accession E-MTAB-209. It was apparent that most differentially methylated regions in HCC samples were hypomethylated (712 probes) in comparison to 168 hypermethylated regions.

Blast2GO and Ingenuity Pathway Analysis (IPA) were performed to characterise the functional relationships between genes with altered methylation in the regions between -1.5kb to +1kb of the TSS in HCC samples compared to control samples. These analyses identified significant networks and functions, such as cell-to-cell signalling, cell cycle, cellular growth and proliferation, cell death, gene expression, DNA replication, cellular assembly and organisation that were enriched in HCC samples compared with control samples. Figures 3.13 and 3.14 show the biological functions associated with genes with altered methylation levels (hyper- and hypo-methylated, respectively) in HCC samples compared to control samples (FDR $< 5\%$). Table 3.1 illustrates a sub-section of the genes (hypo- and hyper-methylated) representing the most relevant biological functions linked to carcinogenesis (the full list of genes and gene ontology analyses are presented in Appendix (File 2, Tables 2 and 3).

Using IPA, cancer was highlighted as the top predicted disorder based on the categories of the genes that were hypomethylated (FDR $< 5\%$, >1.5 -fold change). Figure 3.15 illustrates hypomethylated genes in zebrafish tumours (FDR $< 5\%$, >1.5 -fold change) related to the canonical pathway “molecular mechanisms of cancer” in humans, including genes such as *c-jun*, *shc* and *p21*.

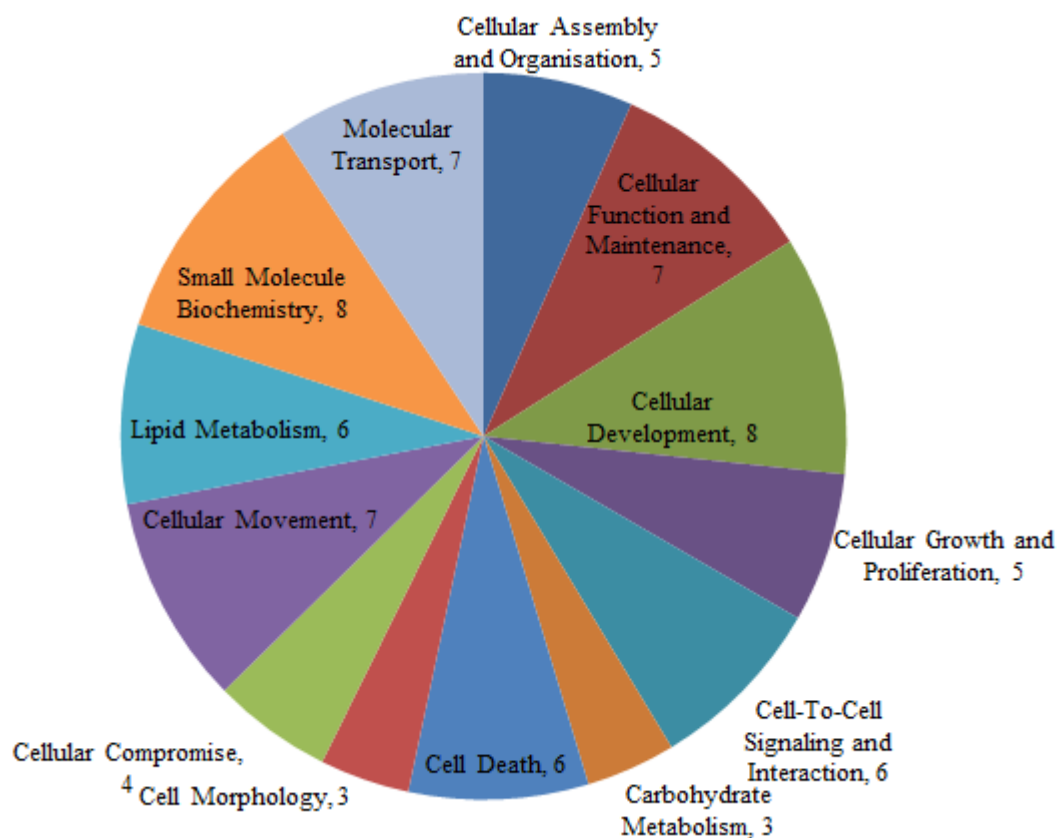


Figure 3.13. Biological functions enriched among genes with hypermethylated levels in HCC samples compared to control samples. IPA was used to group the genes with altered methylation levels based on biological functions. Each section of the pie chart corresponds to the number of the genes associated with the mentioned biological function with altered methylation levels in HCC samples compared to control samples (>1.5-fold change, FDR <5%).

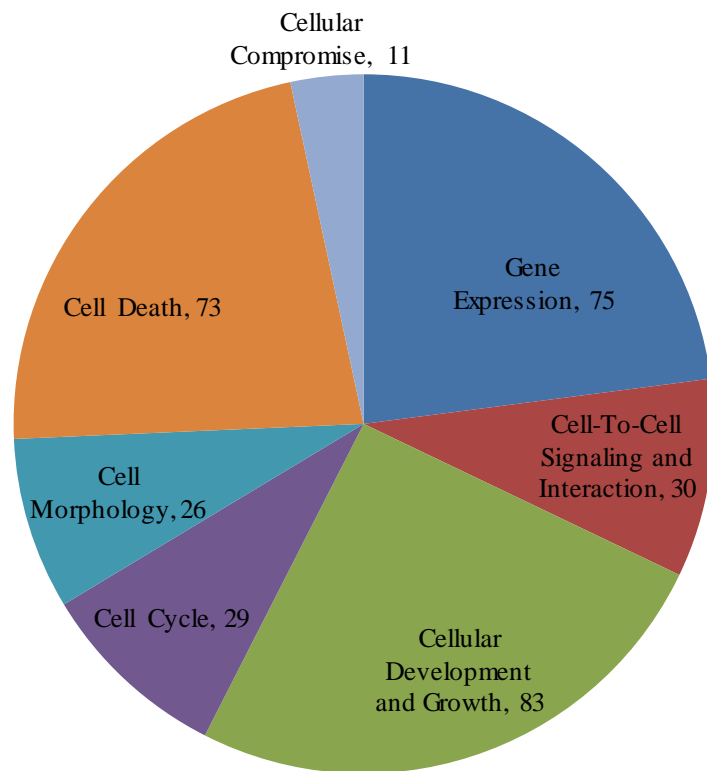


Figure 3.14. Biological functions enriched among genes with hypomethylated levels in HCC samples compared to control samples. IPA was used to group the genes with altered methylation levels based on biological functions. Each section of the pie chart corresponds to the number of the genes associated with the mentioned biological function with altered methylation levels in HCC samples compared to control samples (>1.5 -fold change, FDR $<5\%$).

Gene name	Gene symbol	Gene ID	Chr. no	CGI region
A. Hypomethylated regions proximal to TSS				
Proliferation				
<i>kruppel-like factor 12b</i>	<i>klf12b</i>	ENSDARG000000032197	9	31117162 - 31117483
<i>insulin-like growth factor binding protein 5a</i>	<i>igfbp5a</i>	ENSDARG000000025348	9	49194032 - 49194791
<i>insulin-like growth factor binding protein 2b</i>	<i>igfbp2b</i>	ENSDARG000000031422	9	49290381 - 49290816
<i>insulin-like growth factor binding protein 1b</i>	<i>igfbp1b</i>	ENSDARG000000038666	2	178184 - 178306
<i>insulin-like growth factor-binding protein 2a precursor</i>	<i>igfbp2a</i>	ENSDARG000000052470	6	22745784 - 22746301
<i>estrogen related receptor delta fragment</i>	<i>esrrd</i>	ENSDARG000000015064	18	48225445 - 48225631
<i>Steroid5-alpha-reductase, alpha polypeptide 2a</i>	<i>srd5a2a</i>	ENSDARG000000043587	1	51038733 - 51038845
Stress				
<i>similar to sh2 domain containing 3C</i>	<i>sh2d3ca</i>	ENSDARG000000028099	10	14303115 - 14303263
<i>PI-3-kinase-related smg-1</i>	<i>smg1</i>	ENSDARG000000054570	3	28544220 - 28544532
Glycolysis				
<i>enolase 2</i>	<i>eno2</i>	ENSDARG000000014287	19	4698036 - 4698168
<i>hexokinase 1</i>	<i>hk1</i>	ENSDARG000000039452	13	23684991 - 23685203
Cell cycle, metastasis, adhesion, cell growth, stress				
<i>c-jun</i>	<i>jun</i>	ENSDARG000000043531	20	14274343 - 14274635
<i>bcl2-associated athanogene 5</i>	<i>bag5</i>	ENSDARG000000017316	13	17324708 - 17324877
<i>angiopoietin-like 3</i>	<i>angptl3</i>	ENSDARG000000044365	6	34165085 - 34165237
<i>angiopoietin-1 receptor precursor</i>	<i>tek</i>	ENSDARG000000028663	5	625275 - 625436
<i>ras homolog gene family, member ua</i>	<i>rhousa</i>	ENSDARG000000019709	13	25136705 - 25136882
<i>menage a trois homolog 1</i>	<i>mnat1</i>	ENSDARG000000002077	13	31141636 - 31142019
<i>serine/threonine and tyrosine protein kinase</i>	<i>dstyk</i>	ENSDARG000000000853	22	461194 - 461772

Continued from previous page

DNA binding and regulation of transcription

<i>DNA (cytosine-5-)-methyltransferase 6</i>	<i>dnmt6</i>	ENSDARG00000015566	17	34759411 - 34759539
<i>leucine zipper protein 2 precursor</i>	<i>luzp2</i>	ENSDARG000000068247	18	35625841 - 35625986
<i>histone h2a</i>	<i>h2a</i>	ENSDARG000000001915	1	724191 - 724473
<i>histone deacetylase 4</i>	<i>hdac4</i>	ENSDARG000000041204	9	46375505 - 46375895
<i>homeobox protein hox-b5a</i>	<i>hoxb5a</i>	ENSDARG000000013057	3	20707021 - 20707754
<i>pancreas transcription factor 1 subunit alpha</i>	<i>ptf1a</i>	ENSDARG000000014479	2	27672474 - 27672641
<i>hypothetical protein LOC692291</i>	<i>loc692291</i>	ENSDARG000000012833	17	12689887 - 12690054
<i>metastasis associated 1 family, member 2</i>	<i>Mta2</i>	ENSDARG000000013031	7	17243535 - 17244290
<i>homeobox protein hox-b4a</i>	<i>hoxb4a</i>	ENSDARG000000013533	3	20721064 - 20721722
<i>lysine-specific demethylase 4a</i>	<i>kdm4a</i>	ENSDARG000000018782	6	4575451 - 4575994
<i>rab11 family interacting protein 4 (class II) a</i>	<i>rab11fip4a</i>	ENSDARG000000053855	25	13380260 - 13380369

Gene name	Gene ID	Chr. no	CGI region
-----------	---------	------------	------------

B. Hypermethylated regions proximal to TSS

Anti-angiogenesis

<i>hematopoietically-expressed homeobox protein hhex</i>	<i>hhex</i>	ENSDARG000000074250	12	43558447 - 43558684
--	-------------	---------------------	----	---------------------

Cell-cell adhesion

<i>hypothetical protein LOC678612</i>	<i>loc678612</i>	ENSDARG000000069505	21	27963202 - 27963305
<i>novel protocadherin protein fragment</i>	<i>pcdh7</i>	ENSDARG000000053462	1	55112983 - 55113251

Transporter

<i>ATP-binding cassette, sub-family a, member 5</i>	<i>abca5</i>	ENSDARG000000074041	12	38940033 - 38940251
---	--------------	---------------------	----	---------------------

Continued from previous page

Immune system				
<i>novel protein similar to nuclear factor, IL3 regulated</i>	<i>nfil3</i>	ENSDARG000000071398	22	21400978 - 21401243
<i>novel protein fragment similar to interleukin</i>	<i>il</i>	ENSDARG000000053462	1	55112983 - 55113251
<i>c5a anaphylatoxin chemotactic receptor</i>	<i>c5ar</i>	ENSDARG000000040319	18	45692130 - 45692420
Angiogenesis and oxidative stress protection				
<i>vascular endothelial zinc finger 1</i>	<i>vezf1</i>	ENSDARG000000008247	10	35492470 - 35493037
Membrane				
<i>coronin, actin binding protein 2ba</i>	<i>coro2ba</i>	ENSDARG000000079440	7	33649014 - 33649398

Table 3.1. Subsection of the genes with altered DNA methylation levels detected at -1.5kb to +1kb of the TSS in HCC samples compared to control samples as measured by CGI-tiling microarray. A. Hypomethylated genes (>1.5-fold change, $P<0.05$). B. Hypermethylated genes (>1.5-fold change, $P<0.05$).

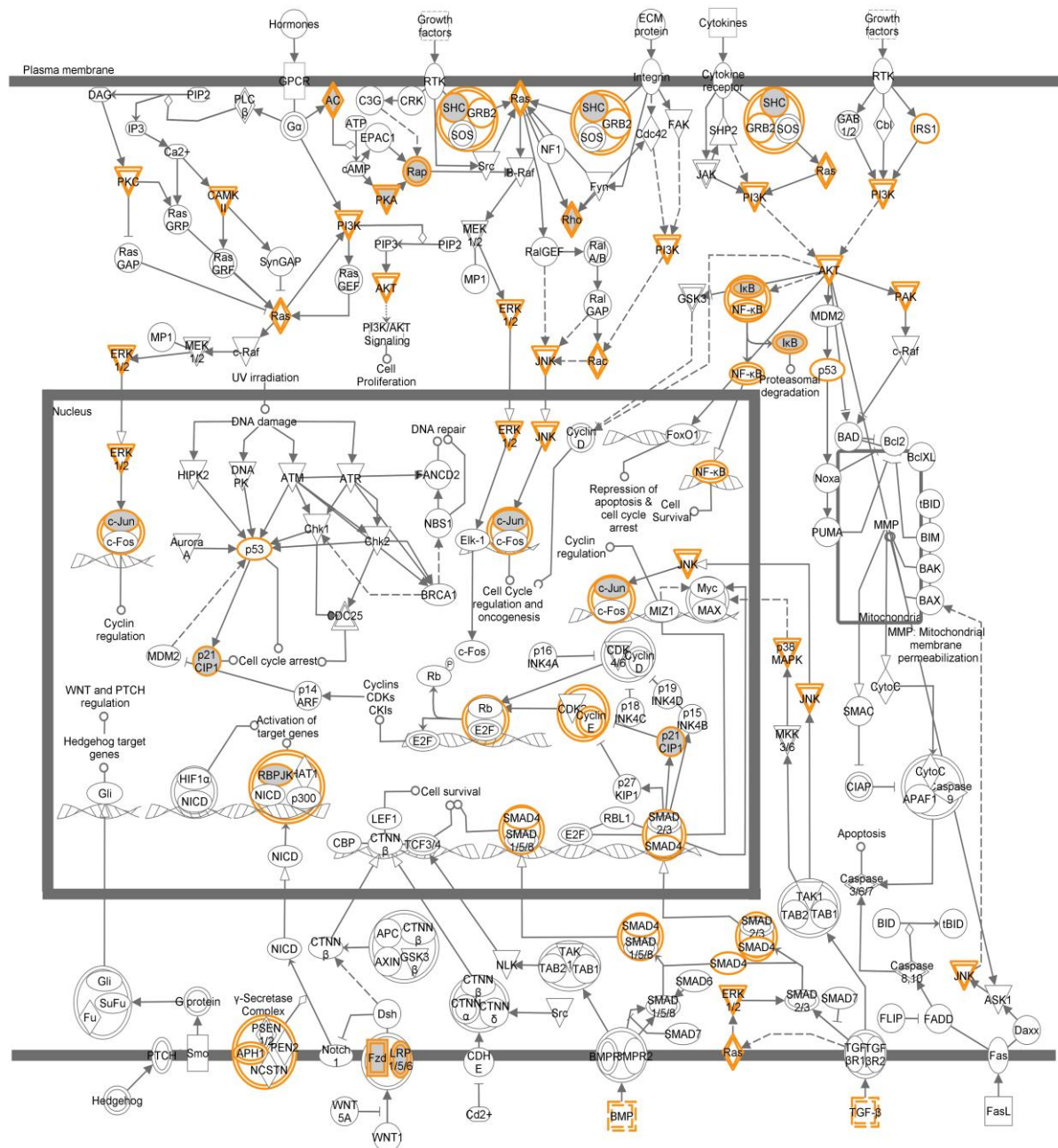


Figure 3.15. Biological network of genes linked to the canonical pathway “molecular mechanisms of cancer” that were hypomethylated (>1.5-fold change, $P < 0.05$) in zebrafish hepatocellular carcinoma compared to healthy liver. This diagram shows the genes that were hypomethylated in zebrafish HCC with grey shading. Additionally, orange outlines indicate the molecules associated with the hypomethylated genes via Ingenuity pathways. Direct interactions are shown as solid lines and indirect as dashed lines. Biological network analysis was performed using Ingenuity Pathway Analysis.

3.3.2.4. Principal components analysis

Using principal components analysis (PCA) scores plot of differentially methylated regions, the four groups of HCC, controls, positive and negative controls were separated based on treatment along the PC1 and PC2 axes (Figure 3.16). PCA is a statistical multivariate data reduction approach used for simplifying complex data sets with several conditions such as microarrays (Raychudhuri et al., 2001). PCA converts a number of correlated variables to a smaller set of uncorrelated variables or principal components (PC), in order to identify patterns within a data set (Yeung and Ruzzo, 2001; Raychudhuri et al., 2001).

The first PC was the main separation factor for the four different categories investigated. In this experiment PC1 accounted for 71.78% of the variance. Negative and positive controls were the two extreme conditions in the experiment in which one was fully un-methylated, while the other was completely methylated. Indeed, this was reflected in the results of the PCA plot, as positive and negative control samples were located at the two ends of the PC1 axis, with controls (healthy tissue samples) and HCC samples located in between.

As shown in the PCA plot (Figure 3.16), healthy zebrafish samples (controls) are clustered closer together compared to tumour samples. This is not unexpected and several factors may contribute to the observed pattern. Control samples are more homogenous in regards to cell type. Although tumours are histopathologically confirmed they are usually a mixture of cells and are likely heterogenous in terms of progression.

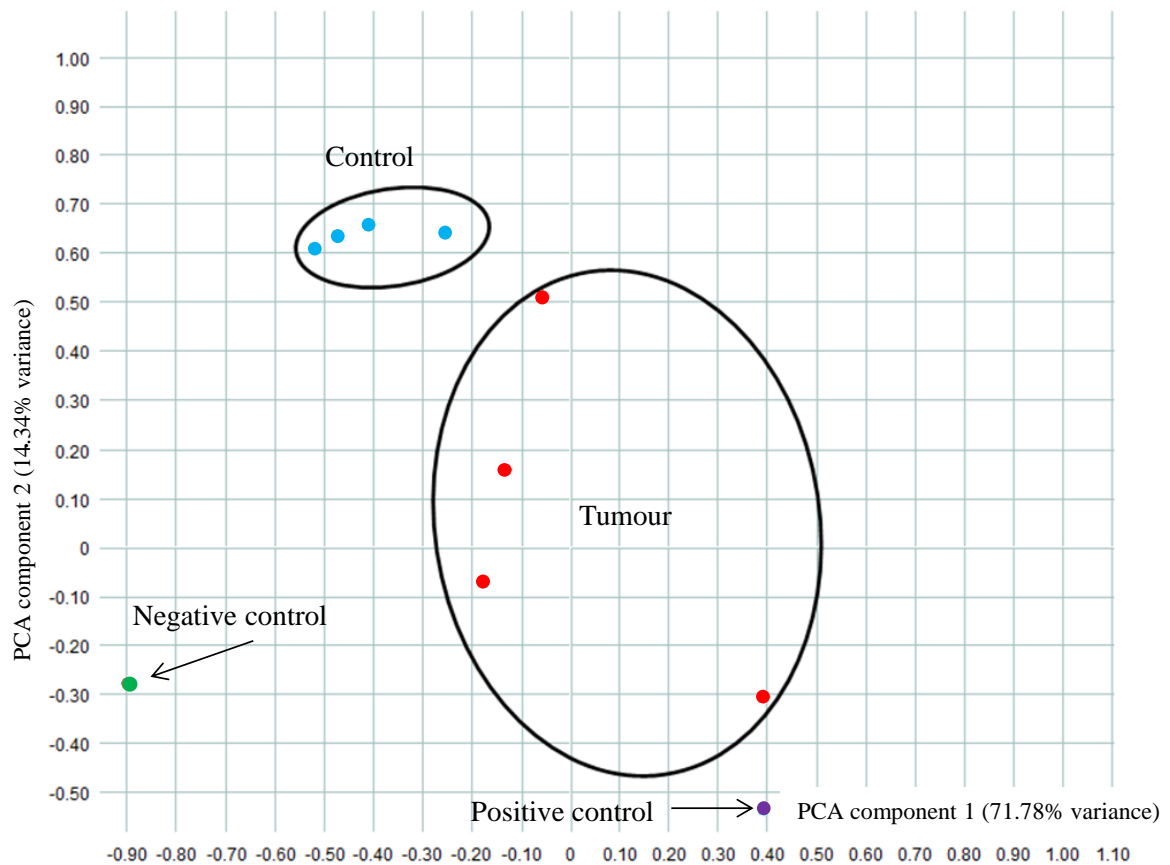


Figure 3.16. Principal component analysis (PCA) scores plot of DNA methylation data. Zebrafish HCC samples (blue), healthy zebrafish liver samples (red), positive control of artificially methylated genomic DNA (purple) and negative control of artificially unmethylated genomic DNA (green) were separated based on treatment along the PC1 and PC2 axes.

3.3.3. Confirmation of the CGI tiling microarray data using bisulfite sequencing PCR

Based on the data obtained from the DNA methylation tiling microarray, three genes from the 2 categories of hypomethylated and hypermethylated genes in HCC samples compared to control samples were selected for validation of the microarray data by bisulfite sequencing PCR (BSP) (three tumour samples and three control samples). The genes selected for additional analysis with BSP were *coronin*, *actin binding protein 2ba* (*coro2ba*) (hypermethylated in HCC), *insulin-like growth factor binding protein 1b* (*igfbp1b*) (hypomethylated in HCC) and *angiopoietin-like 3* (*angptl3*) (hypomethylated in HCC). The primers and annealing temperatures used are shown in Chapter 2, section 2.6.3, Table 2.2. Measurement of peak areas for C and T bases at a particular CpG site and calculation of the percentage of DNA methylation in four-dye trace sequencing data, allowed a semi-quantitative assessment of the DNA methylation level at the specified site. To compare the methylation status of the targeted area by BSP in tumour and control, the total amount of methylation for the region was measured and compared. BSP results for *igfbp1b*, *coro2ba* and *angptl3* confirmed the changes observed from the tiling microarray (Figures 3.17 and 3.18).

The CGI of the *gstp1* gene, as shown in Figure 3.5, is un-methylated in both tumour and control samples. Following treatment with CpG methyltransferase, *SssI*, in the presence of SAM, it became fully methylated (Figure 3.9). Therefore, this artificially methylated gene was used as a positive control in BSP experiments to confirm the semi-quantitative measurements of DNA methylation.

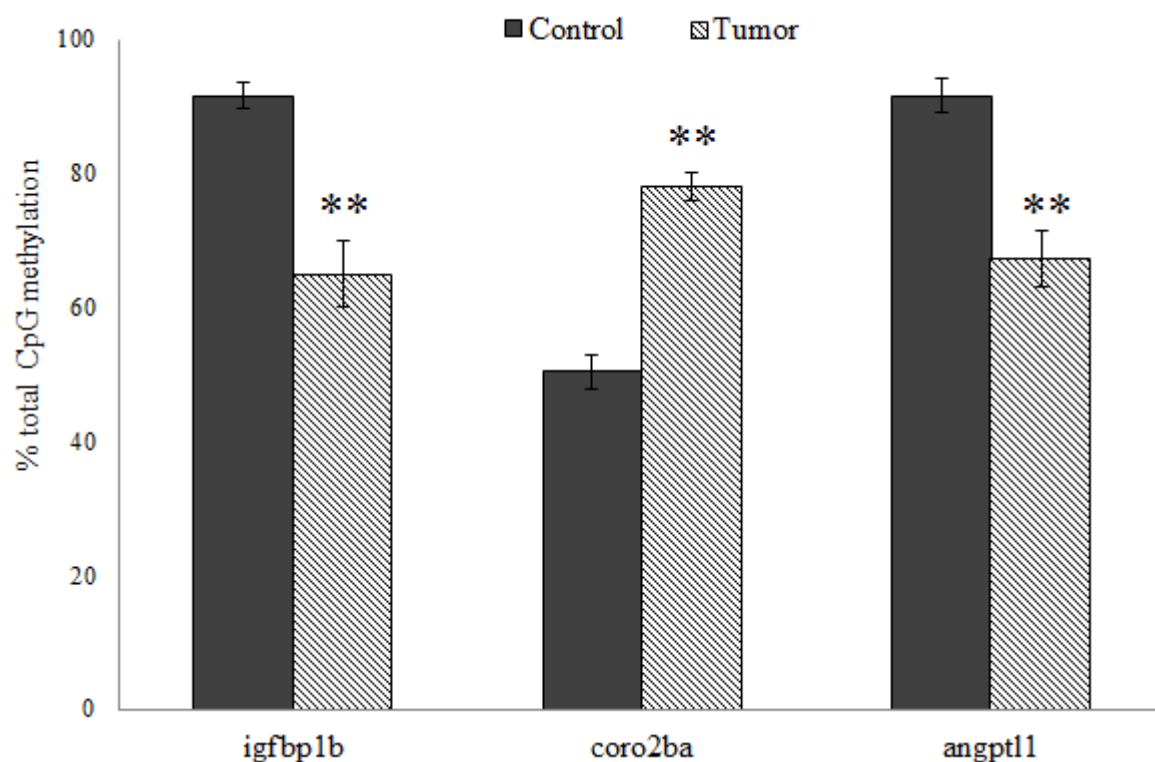


Figure 3.17. Confirmation of the CGI tiling microarray using bisulfite sequencing PCR. Percentage of DNA methylation (combined for all measured CpG dinucleotides) in *igfbp1b*, *coro2ba* and *angptl1* genes showed significant differences between tumour and control samples (** $P < 0.01$). (*igfbp1b*: number of CpG sites in tumours = $15 \pm \text{SEM}$, number of CpG sites in controls = $15 \pm \text{SEM}$; *coro2ba*: number of CpG sites in tumours = $51 \pm \text{SEM}$, number of CpG sites in controls = $51 \pm \text{SEM}$; *angptl1*: number of CpG sites in tumours = $21 \pm \text{SEM}$, number of CpG sites in controls = $21 \pm \text{SEM}$). Note: number of CpG sites is the sum of all CpG sites investigated within and between each replicates. P -value determined by 2-tailed Student's t -test.

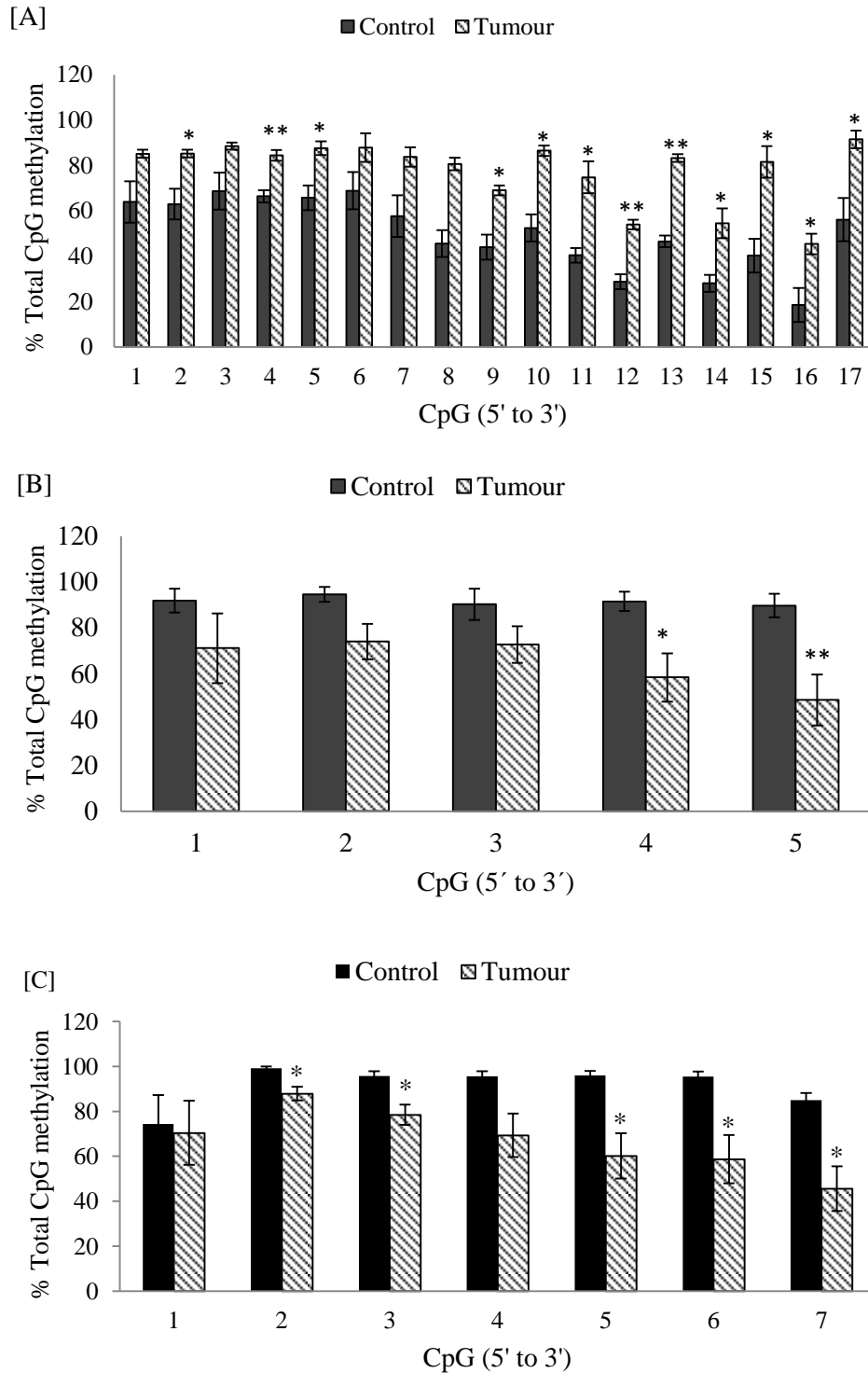


Figure 3.18. Methylation levels in tumour and control samples for individual CpG sites. A. *coronin, actin binding protein 2ba (coro2ba)* CGI region: -879 to -580, B. *insulin-like growth factor binding protein 1b (igfbp1b)* CGI region: -488 to -259 C. *angiopoietin-like 3 (angptl3)* CGI region: +106 to +258. n= 3 independent controls and 3 independent tumours \pm SEM. * $p < 0.05$, ** $p < 0.01$, p -value determined by 2-tailed Student's t -test (tumour was compared to control for each CpG site). The probes and primers were designed based on version 56, genome build zv8, genome build date April 2009.

3.3.4. Gene expression and DNA methylation

In a study published by Lam et al., who also provided the zebrafish DNA samples (Lam et al., 2006), transcriptional changes in zebrafish HCC samples compared to control adult zebrafish liver samples were established. In addition, they compared the statistically significantly altered transcripts in zebrafish HCC samples with four different types of tumours found in humans (liver, gastric, prostate and lung cancer). Their findings demonstrated that hallmarks of cancer in humans were also altered in the zebrafish tumours (e.g. genes associated with cell cycle and proliferation, apoptosis, DNA replication and repair, and metastasis). However, gene expression signatures in zebrafish liver tumours most significantly correlated with gene expression profiles in liver tumours in humans (e.g. changes in Wnt- β -catenin and RAS-MAP kinase pathways) (This paper is provided as Appendix File 4).

The data from this study was used to compare the alterations in gene expression with alterations in DNA methylation in HCC samples compared to healthy zebrafish livers. Comparisons are limited to the genes represented on the expression microarray that were differentially expressed and genes represented on the CGI tiling microarray. Therefore, DNA methylation and expression could only be compared for 194 genes. From these 194 genes with altered gene expression, 68 genes had CGIs 1.5 kb upstream to 1kb downstream of their TSS. Due to the criteria used for the design of the probes, only 49 genes from the previously identified 68 genes were represented on the tiling microarray and comparison were made between DNA methylation levels and gene expression. In total, 22 genes were identified with both significantly altered transcription and methylation (>1.5-fold change) in zebrafish HCC samples (Table 3.2, A full list of the 194 genes investigated is presented in Appendix File 5). Most of the genes identified were hypomethylated with up-regulated gene expression, such as *mitogen-activated protein kinase 1 (mapk1)*, *cell division protein kinase 8 (cdk8)*, *RAB2A*, *member ras oncogene family (rab2)* and *proliferating cell nuclear antigen (pcna)*. This is not un-expected as most of the genes identified in our experiment were hypomethylated. Therefore, an extensive proportion of the 22 genes were comprised of hypomethylated genes (from the 29 probes (representing 22 genes) significantly altered in DNA methylation (>1.5-fold change, $p<0.05$) 20 probes were hypomethylated and 9 probes were hypermethylated). IPA was applied to the list of genes that were significantly hypomethylated with up-regulated gene expression. The network highlighted in Figure 3.19 contains genes involved in formation

of cancer. For example, *pcna* is involved in formation of the replication fork and directing maintenance methyltransferases to the newly synthesised DNA strand. *mapk1*, *capn2* and *erk1/2* were a few more genes identified in this network with a known role in tumourigenesis.

Gene name	Symbol	Chr. no	Expression level	Methylation level	Gene ID
<i>matrix metalloproteinase 14 (membrane-inserted) alpha</i>	<i>mmp14a</i>	7	↑	↓	ENSDARG00000002235
<i>histidyl-tRNA synthetase</i>	<i>hars</i>	14	↑	↓	ENSDARG00000003693
<i>tubulin, alpha 8 like 4</i>	<i>tuba8l4</i>	6	↑	↓	ENSDARG00000006260
<i>e2f transcription factor 6</i>	<i>e2f6</i>	20	↓	↓	ENSDARG00000008119
<i>acetoacetyl-coa synthetase</i>	<i>aacs</i>	5	↑	↑	ENSDARG00000012468
<i>nucleophosmin 1</i>	<i>npm1</i>	10	↑	↓	ENSDARG00000014329
<i>cyclin dependent kinase 8 (probe 54)</i>	<i>cdk8</i>	24	↑	↓	ENSDARG00000016496
<i>cyclin dependent kinase 8 (probes 21, 31)</i>	<i>cdk8</i>	24	↑	↑	ENSDARG00000016496
<i>hepatoma-derived growth factor-related protein 2</i>	<i>hdgfrp2</i>	22	↑	↓	ENSDARG00000019530
<i>rab2a, member ras oncogene family</i>	<i>rab2</i>	2	↑	↓	ENSDARG00000020261
<i>lim domain containing preferred translocation partner in lipoma</i>	<i>lpp</i>	6	↑	↓	ENSDARG00000023578
<i>Insulin-like growth factor binding protein 5a</i>	<i>igfbp5b</i>	9	↓	↓	ENSDARG00000025348
<i>mitogen-activated protein kinase 1</i>	<i>mapk1</i>	5	↑	↓	ENSDARG00000027552
<i>cancer susceptibility candidate gene 3 protein homolog</i>	<i>casc3</i>	3	↑	↑	ENSDARG00000029911

Continued from previous page					
<i>hnrpa0l protein</i>	<i>hnrpa0l</i>	14	↑	↓	ENSDARG00000036161
<i>cyclin T2</i>	<i>ccnt2</i>	9	↑	↓	ENSDARG00000036685
<i>proliferating cell nuclear antigen</i>	<i>pcna</i>	10	↑	↓	ENSDARG00000054155
<i>ret proto-oncogene (probes 16, 19)</i>	<i>ret</i>	13	↓	↑	ENSDARG00000055305
<i>ret proto-oncogene (probes 20, 24)</i>	<i>ret</i>	13	↓	↓	ENSDARG00000055305
<i>integrin-linked kinase (probes 8, 1)</i>	<i>ilk</i>	10	↑	↓	ENSDARG00000056964
<i>integrin-linked kinase (probes 6)</i>	<i>ilk</i>	10	↑	↑	ENSDARG00000056964
<i>novel protein similar to vertebrate threonyl-tRNA synthetase</i>	<i>tars</i>	18	↑	↑	ENSDARG00000075429
<i>ribonuclease inhibitor 1</i>	<i>rnh1</i>	22	↓	↓	ENSDARG00000078234
<i>calpain 2, (m/II) large subunit, like</i>	<i>capn2</i>	22	↑	↓	ENSDARG00000034211
<i>synaptopodin-2</i>	<i>synp2</i>	7	↓	↑	ENSDARG00000079675

Table 3.2. Genes with both significantly altered gene expression levels and DNA methylation levels (>1.5-fold change, $p < 0.05$) (expression level: ↑ up regulated, ↓ down regulated; methylation level: ↑ hypermethylated, ↓ hypomethylated).

3.4. Discussion

The impact of DNA methylation upon the regulation of gene expression and its implications in disease has been studied extensively in relation to human health. As a result, the links between environmental agents such as chemicals or internal factors such as hormonal imbalances, alterations in epigenome and diseases in rodents and humans have become more apparent (discussed in Chapter 4). These findings have highlighted the importance of considering epigenetic mechanisms in ecotoxicology (Vandeghechuchte and Janssen, 2011) and aquatic diseases. Therefore, the main aim of this chapter was to provide a starting point and a source of information for studying DNA methylation in fish tumours.

3.4.1. Overall DNA methylation levels in fish

Overall DNA methylation levels vary throughout the animal kingdom with the highest levels reported in vertebrates, intermediate levels in non-arthropod invertebrates and the lowest levels of DNA methylation in arthropods (Bird, 1980). However, there are some exceptions, for example DNA methylation has not been detected in the non-arthropod invertebrate, the nematode *Caenorhabditis elegans* (Simpson et al., 1986). In addition to DNA methylation levels, DNA methylation patterns differ throughout the animal kingdom. In invertebrate genomes, regions with high levels of methylation are separated by equally large regions of unmethylated DNA. This gives rise to a pattern referred to as mosaic pattern of DNA methylation. In contrast, the genomes of vertebrates have a more globally homogenous pattern of methylation. In this case methylated DNA regions are distributed evenly throughout most of the genome. This leaves the CGIs located at the promoter regions of the genes as the regions that are free from methylation (Bird, 2002; Pogribny, 2010). Vertebrates are further divided into two distinct populations based on overall CpG methylation levels. The genomes of homeothermic vertebrates have almost two-fold lower methylation levels compared to ectothermic vertebrates (Varriale and Bernardi, 2006a). For example 4_5% of all the cytosines are methylated in the DNA extracted from the human liver and *Bos taurus* thymus while this value reaches around 7_10% in the DNA extracted from the liver of fish (Aniagu et al., 2008). This is in agreement with our findings in which the DNA methylation levels of the three fish species (~8%) investigated were two-fold higher than the DNA extracted from the thymus of mammalian species, *Bos taurus* (~4%). Several theories have been proposed for the

observed differences in overall DNA methylation levels. Bird (1995) proposed that the relatively high levels of DNA methylation in vertebrates is an evolutionary strategy evolved to decrease the background transcriptional noise due to increased number of genes in vertebrates compared with invertebrates (i.e. increased number/multifunction of genes and their products that are essential for development of their complex body). This theory explains the differences observed in DNA methylation levels and patterns between vertebrate and invertebrate genomes. However as humans compared to fish generally have a more complex genome (in regards to regulation, differentiation, interactions and functions), this theory does not account for the higher levels of DNA methylation observed in ectotherm genomes compared to homeotherm genomes.

A second set of studies suggest that the changes in overall DNA methylation levels in vertebrates inversely relate to body temperature (i.e. methylation level: homeothermic vertebrates < tropical fish < temperate fish < polar fish) (Varriable and Bernardi, 2006a; Varriable and Bernardi, 2006b; Jabbari and Bernardi, 2004). Increase in body temperature results in higher rate of deamination of 5-methylcytosine. This causes an increase in 5mCpG:TpG transition in homeothermic species, leading to lower levels of DNA methylation in homeothermic animals compared to ectothermic animals. Furthermore, transition of 5mCpG to TpG in homeothermic vertebrate genomes is accompanied by accumulation and over-representation of the complementary dinucleotides TpG and CpA in their genome (Varriable and Bernardi, 2006a; Bird, 1980; Spontaneous mutation of 5mC is discussed in Chapter 1, section 1.3.5.1.3). However, the reason for the observed differences in overall DNA methylation levels remains controversial. Nevertheless, what is in agreement is that these differences observed in DNA methylation patterns, levels and sites throughout the animal kingdom are indicative of the possibility that the function of DNA methylation has altered as species have evolved (Lee et al, 2010). Therefore, to understand the role of DNA methylation it is important to establish the pattern of DNA methylation (Bird, 2002).

3.4.2. DNA Methylation alterations at gene level in zebrafish HCC samples

A range of carcinogenic polycyclic aromatic hydrocarbons (PAHs) including DMBA can initiate DNA methylation changes, including global DNA hypomethylation. Therefore in addition to inducing mutations, these chemicals can initiate and contribute to carcinogenesis

process by causing heritable changes in DNA methylation patterns (Wilson et al., 1987). In this study, treatment with DMBA was conducted for a short period of time. Following chemical exposures, fish were kept for an extensive period of time (6_10 months) in chemical-free water. A number of methylation changes are formed as a result of the initial chemical exposure (i.e. initiating events). These epigenetic events are maintained as a memory of the exposure to the chemical. However, most likely the high proportion of the observed changes in DNA methylation profiles are characteristic of the type of formed tumour. This is due to the extensive chemical-free period. Nevertheless, studies have demonstrated that DNA methylation profiles depend on both the causative agents (memory of the exposure) (Pei et al., 2009; Prins, 2008) and the type of tumour (Esteller and Herman, 2002). In this study DNA methylation modifications were investigated in relation to the features associated with tumours rather than the causative agent by using a custom-made promoter CGI zebrafish microarray. On this microarray, each CGI was represented by a number of probes covering the entire length of the CGI. This provided a comprehensive coverage of the CGI. The analyses indicated that DNA methylation levels are not uniformly distributed across the regions investigated in the healthy adult zebrafish liver samples. Indeed, genes with specific biological functions were enriched or under-methylated compared to the average DNA methylation levels of all the investigated regions (Results, section 3.3.2.2). As discussed in Chapter 1, DNA methylation is involved in regulation of transcription and tissue-specific transcription. Therefore, the changes in methylation levels could reflect the differences in transcription levels of these genes. Although all somatic cells contain the same genetic material, they are phenotypically and functionally diverse. This, in part, has been linked to tissue-specific differentially methylated regions which consequently will alter the transcription levels of these regions, providing each tissue with a unique phenotype (Christensen et al., 2009; Eckhardt et al., 2006). However, this theory requires testing and comparing the achieved DNA methylation profile with DNA methylation profiles of other adult zebrafish tissues. Therefore at this stage the established normal DNA methylation profile can only be used as a measure of background level of DNA methylation for comparison with treatment conditions in adult zebrafish liver (i.e. relative changes in DNA methylation levels following exposure to chemicals or liver diseases). Therefore DNA methylation profiling of healthy adult zebrafish liver has provided: 1. A list of genes with CpG islands at their promoter regions 2. A list of genes with different levels of DNA

methylation at their promoter regions (relative levels rather than absolute levels of DNA methylation). The identified genes can be used in any future studies as suitable target genes to test whether their DNA methylation is altered as a response to treatment or whether DNA methylation may have a role in regulation of their transcription levels. This bypasses the initial screening of the genes that was conducted at the beginning of this study for the identification of methylated genes in zebrafish (see results, section 3.3.2.1). Thus although this method provided a relative measurement of changes in DNA methylation of HCC tissues compared to healthy tissues and relative measurements of the levels of enrichment of methylated regions rather than absolute levels of DNA methylation in control samples, this did not influence the identification of the categories of genes that were enriched/under-represented in the analysed groups. Therefore, the data achieved from the MeDIP-tiling microarray provided a useful insight into alterations to DNA methylation related to several biological functions associated with tumourigenesis, such as regulation of transcription in HCC samples.

Regulation of transcription is a complex procedure. It is modulated and controlled at different levels and through various mechanisms. Epigenetic mechanisms are key elements in this process with histone-DNA interactions in nucleosomes, covalent modification of the N-terminal tails of histones, RNA interference (RNAi) and DNA methylation being some of the main epigenetic regulatory mechanisms (Momparker and Bovenzi, 2000; Zilberman et al., 2007; Cui et al., 2010). As shown in Table 3.1A, GO terms associated with DNA binding and transcription regulation were significantly over-represented amongst genes hypomethylated in HCC samples. For example, methylation of the *histone H2A* gene was reduced in the HCC samples compared to controls. As hypomethylation of genes has been associated with increased expression (Lopez et al., 2009; Rauch et al., 2007), it is possible that this change could cause increased expression of *histone H2A* gene. Increase in expression of this gene could further affect the composition of nucleosomes and thereby regulation of transcription (Turner, 2007). Indeed, histones can cover the transcription start sites and transcription termination sites (TTS) of the genes, restricting access of the transcription factors to these regions and resulting in altered transcription levels (Cui et al., 2010). In addition to the role of globular domains of the histone in regulation of transcription, post-transcriptional chemical modifications of the N-terminal tail of histones are extremely important in regulation of gene expression and replication (Turner, 2007; Zhou et al., 2000; Turner, 2009). Indeed, the

methylation of *histone deacetylase 4* and *lysine-specific demethylase 4A* were reduced in the HCC samples compared to controls. The N-terminal acetylation or deacetylation of histones, mediated by histone deacetylases and histone acetyltransferases, can decrease and increase the positive charge of the histones, respectively, and as a result affect the interaction between DNA and histones. This subsequently can permit or restrict the access of transcription factors to the regulatory regions of the genes (Zhou et al., 2000; Turner, 2007; Lopez et al., 2009; Turner, 2009).

Methylation profiles of DNA in cancer cells are extensively distorted (Lopez et al., 2009), thus mediating changes in most of the key pathways (e.g. cell cycle, apoptosis, proliferation) associated with tumourigenesis (Baylin et al., 2001). This is in agreement with our findings using Ingenuity Pathway Analysis (IPA). This analysis indicated that the genes with altered methylation in zebrafish hepatocellular carcinoma were associated with biological functions such as cell death, cell morphology, inflammatory response, DNA repair and replication. The changes in methylation levels of these particular genes and pathways could be directly or indirectly linked to their altered expression levels during tumourigenesis.

Most importantly, GO terms associated with genes involved in cell proliferation and regulation of cell death were significantly over-represented in the list of hypomethylated genes in zebrafish HCC samples. For example our results showed a significant decrease in the DNA methylation of a positive regulator of the B-cell lymphoma 2 (Bcl-2) protein, *bcl-2 associated athanogene 5 (bag5)* (Yang et al., 2007). It is known that anti-apoptotic genes, such as *BCL-2*, and their regulators, are often over-expressed in human tumours (Yang et al., 2007; Yang et al., 1997). This finding is particularly relevant since an imbalance between regulation of cell proliferation pathways and cell death by apoptosis can promote the development of tumours (Sakinah et al., 2007; Klaunig et al., 2000).

In addition to the observed changes in DNA methylation levels of genes involved in anti-apoptotic pathways in zebrafish HCC, DNA methylation levels of genes involved in proliferation pathways were also altered. Insulin like growth factors (IGF) and insulin like growth factor binding proteins (IGFBPs) play important roles in organising cell proliferation, apoptosis and differentiation and are commonly deregulated in human tumours (Inman et al., 2005). In zebrafish HCC samples, genes for several insulin growth factor binding proteins such as *igfbp2b* were significantly hypomethylated. The promoter region of the human

IGFBP-2 gene is rich in CpGs and lacks a TATA box (Hoeflich et al., 2001). It is therefore plausible that methylation plays an important role in regulating the expression of this gene. Multiple complex IGF-dependent and independent biological functions influenced by the tissue type and pathological status have been identified for IGFBPs (Inman et al., 2005; Hoeflich et al., 2001). An increased level of IGFBP-2 protein has been reported in liver tissues and serum during human malignancy (Inman et al., 2005; Hoeflich et al., 2001) with a positive correlation to the malignancy status of the tumour (Hoeflich et al., 2001). In contrast to its normal role as a negative regulator of growth, increased levels of IGFBPs in tumours have been linked to enhanced proliferation, partly as a response to androgens and hypoxia-inducible factor-1 (HIF-1) protein. Thus, anaerobic conditions can result in increased levels of IGFBPs (Kelly et al., 2008).

A lack of vascular supply at the early stages of tumourigenesis in highly proliferating tumours results in hypoxia (Kelly et al., 2008). Under hypoxic circumstances glycolysis becomes the dominant pathway for energy production in tumours and glycolytic enzymes are induced. In association with this, the HIF-1 protein, expressed in an anaerobic environment, initiates the transcription of several genes involved in stress and glycolysis as well as *IGFBP2* (Hoeflich et al., 2001; Kelly et al., 2008). This is in agreement with the findings from this study in which the GO terms related to glycolysis and hypoxia pathways were more prevalent in the list of hypomethylated genes in HCC samples. Glycolytic enzymes such as enolase 2 (*ENO2*) and hexokinase 2 showed significant decreases in methylation of their genes in HCC samples, indicating a potential increase in expression. Increased expression of enolases such as *ENO1* and 2, as a response to hypoxia and HIF-1, has been reported in human HCC (Hamaguchi et al., 2008). As well as its function in energy production, *ENO1* has been associated with enhanced proliferation in HCC (Hamaguchi et al., 2008). Therefore, based on the functions of the genes where methylation was significantly altered in HCC samples, it appears that there is a link between induction of IGF, IGFBPs, HIF, anti-apoptotic and glycolytic pathways (Hoeflich et al., 2001; Kelly et al., 2008; Feldser et al., 1999). This is similar to the findings previously reported on gene expression in human HCC, implying that differential DNA methylation is at least partially causative of differential gene expression in HCC. However, it has to be noted that some of the observed differences in DNA methylation of the genes in HCC samples compared to control samples can be caused by change in the expression levels of these genes (this concept is further discussed in Chapter 6).

3.4.3. Confirmation of the MeDIP-tiling microarray data using Bisulfite sequencing PCR

Direct sequencing of bisulfite-treated and amplified DNA was used for the measurement of average methylation percentage in a population of DNA molecules. Both normal and tumour tissues are heterogeneous in terms of molecular alterations and cell populations. As in this study DNA was extracted from a tissue with a mixture of cell populations, complete homogeneity of the sequenced data was not anticipated. Therefore, the proportion of C/T was compared between samples based on well-established methodology previously utilised (Mill et al., 2006; Lewin et al., 2004; Paul and Clark, 1996). Hence this method (as well as allowing the detection of partial and rare events (Rein et al., 1998)) provided an average percentage of DNA methylation at particular CpG sites in a population of DNA molecules. This is in contrast to establishment of the methylation status of a CpG in one molecule achieved by cloning prior to sequencing (Mill et al., 2006; Lewin et al., 2004; Rein et al., 1998; Paul and Clark, 1996). However, it has to be noted that although BSP is the gold-standard method for measurement of DNA methylation levels with single base resolution (Suzuki and Bird 2008), this method cannot distinguish between 5-mC and 5-hydroxymethylcytosine (5-hmC). The latter reacts with bisulfite and yields cytosine 5-methylenesulfonate (CMS). CMS, similar to 5-mC, does not readily deaminate and therefore it is detected as a cytosine base in the bisulfite treated, sequenced data (Huang et al., 2010). Nevertheless, this method has remained as the most reliable method for measurement of DNA methylation levels. High density of 5-hmC interferes with the activity of DNA polymerase and sequencing of the bisulfite treated DNA (Huang et al., 2010). Therefore as no complication was encountered during amplification step and due to agreement of BSP data with MeDIP-tiling microarray data, the presence of 5-hmC in the investigated genes with BSP seems less likely.

3.4.4. Comparison of gene expression and DNA methylation data

As distortion of DNA methylation profiles of cells can give rise to altered expression of genes commonly implicated in cancers (Momparler and Bovenzi, 2000), the identified genes with significantly altered DNA methylation levels in HCC samples were compared with the differentially expressed genes from the study of Lam et al., (2006).

Although comparison of the genes with statistically significantly altered DNA methylation and expression levels were limited to the genes present on both microarrays (49 genes), 22 genes showed both altered DNA methylation and transcription. These genes were associated with biological functions that are commonly implicated in cancers, such as proliferation, apoptosis, differentiation, migration, and transcription regulation. For example both *cyclin T2* and *cyclin dependent kinase8 (cdk8)*, involved in transcription regulation and cell-cycle, were hypomethylated and over-transcribed in HCC samples (Yang et al., 2004; Firestein et al., 2008). Most notably, half of the hypomethylated genes over-transcribed in HCC samples were identified to be directly or indirectly involved in detachment of cancer cells from surrounding tissues and metastasis, such as *matrix metalloproteinase14 (mmp14)*, *lim domain containing preferred translocation partner in lipoma* and *calpain2*. Metastasis and invasion is one of the essential alterations in the physiology of cancer cells that allows growth and spreading of the tumour cells (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011).

MMP14 is involved in activation of MMP2. MMPs are proteolytic enzymes over-expressed in many cancers (Clark et al., 2008) and have been found to be activated through hypomethylation in pancreatic cancer (Sato et al., 2003). Due to their protein degradation activity and breakdown of extracellular matrix, they are important in tumour invasion and angiogenesis. Furthermore, their ability to detach IGF from the IGFBPs leads to promotion of growth in tumours (Egeblad and Werb, 2002). In addition, calpains, including calpain2, which are calcium activated endopeptidases, are also found to be over-expressed in cancers and have been associated with invasion of cancer cells as well as apoptosis and proliferation (Mamoune et al., 2003; Nakagawa and Yuan, 2000). It has been discovered that over-expression of calpain2 in prostate cancer is regulated through epigenetic mechanisms (Mamoune et al., 2003), which agrees with hypomethylation of the calpain2 promoter in this study.

Lam et al., (2006) reported that expression of Wnt/ β -catenin and Ras-MAPK (mitogen-activated protein kinase) pathways were altered in zebrafish HCC samples. The data from this study also showed that genes involved in these two pathways were altered at the methylation level in these samples. *erk1* involved in the ERK pathway (extracellular-signal-regulated kinases), an evolutionary conserved MAPK pathway, was hypomethylated and over-transcribed in HCC samples. From the Wnt/ β -catenin pathway, *frizzled homologs 8a* and *8b*, *secreted frizzled-related protein 2 (sfrp2)* and *c-jun*, a target gene for β -catenin and ERK

pathway, were hypomethylated in zebrafish HCC samples. Both pathways are associated with proliferation, apoptosis, differentiation and migration in cancers and are commonly deregulated in human liver cancers (Dhillon et al., 2007; Firestein et al., 2008; Bikkavilli and Malbon, 2009; Lam et al., 2006).

However, regulation of gene expression through CGI methylation is extremely complex. For example, the data presented in this chapter (Table 3.2) demonstrated that methylation and expression levels of several genes were altered in the same direction (e.g. *cancer susceptibility candidate gene 3 protein homolog* was hypermethylated and over-transcribed in HCC samples) or in the case of suppressed *ret proto oncogene*, two regions of the CGI were hypomethylated while two other sections were hypermethylated. As discussed (Chapter 1, section 1.3.2.3) the relationship between gene expression and DNA methylation is complex and several factors affect the relationship between them.

Although the methylation level of an entire CGI located around the TSS of a gene could be altered, the methylations of particular CpG sites known as core regions are of key importance in determining the effects of methylation upon transcription. These core regions could be as short as single CpG dinucleotides. For example, methylation of a single CpG site in the TSS region of *tristetraprolin* gene (*TTP*), a negative post transcriptional regulator of c-MYC, results in transcriptional silencing of this gene in liver cancers (Van Vlodrop, 2011). In addition, the presence of transcriptional enhancers and suppressors, alternative promoters at intra- and intergenic CpG dinucleotides, regulation of transcription of several genes by one promoter, activity of more distal response elements and location of several CpGs outside the traditionally defined CGIs can influence the outcome in terms of gene expression and thus interpretation of the data (Siegfried and Simon, 2010; Jones, 1999; Van Vlodrop, 2011). Therefore establishing a relationship between gene expression and DNA methylation data without structural and mechanistic analysis and identification of core CpG regions is not entirely possible.

3.4.5. Comparison between DNA methylation profiles of HCC and healthy liver in zebrafish and in human

Lam et al., (2006) demonstrated that the changes in the expression of key genes involved in promotion of tumours are conserved in the phylogenetically distant species, zebrafish and

human. This emphasised the importance of these genes in cellular responses and strengthened the suitability of fish as models for carcinogenesis research. To further investigate the possibility of conservation of DNA methylation profiles between zebrafish and human healthy livers and HCC tumours, the data achieved from this study were compared with a study by Archer et al., (2010), focusing on DNA methylation profiling of human HCC tumours. This study was selected based on their experimental approach, comprehensive analysis and design. In this study an Illumina Golden Gate methylation bead microarray was used to investigate the methylation of 1,505 CGIs (807 genes) in 76 patients (20 hepatitis C virus (HCV) induced HCC tissues, 20 adjacent non-tumourous samples from HCC containing livers, 16 HCV cirrhotic tissues from patients without HCC, 20 normal liver tissues). As well as high numbers of genes and patients investigated in this study, the effect of several factors in their experimental design was considered, such as examination of both surrounding tissues and liver samples from healthy individuals (DNA methylation differences in surrounding tissue and non-tumour containing tissue is covered in Chapter 4), analysis of samples with similar age groups and most importantly, HCC samples with same underlying cirrhosis etiology (HCV induced HCC) were used in their experiment. The latter is important as it has been shown that HCV induced HCC and hepatitis B virus (HBV) induced HCC samples have different methylation profiles (Pei et al., 2009).

In this study, non-tumour containing livers from zebrafish and HCC tumour samples were used. In order to match the experimental conditions as best as possible between the human DNA methylation study conducted by Archer et al., (2010) and the zebrafish study presented in this chapter, comparisons between human and zebrafish methylation profiles were restricted to human DNA methylation data obtained from healthy liver and HCC samples from humans. Therefore in these comparisons the human data obtained from adjacent non-tumourous samples from HCC containing livers were not used. A low correlation was detected between the methylation pattern of hepatic genes in healthy human liver presented by Archer et al. (2010) and healthy zebrafish liver in this study ($r=0.187$). A poor correlation was also identified between HCV induced liver tumours in humans and HCC tumours in zebrafish ($r=0.166$). This implied that the methylation patterns between human and zebrafish at individual gene levels in both healthy liver and HCC samples are not conserved (The list of genes common between the two studies are presented in Appendix File 6). This could be partly explained due the nature of the zebrafish genome. Zebrafish often have more members

of multigene families than humans (Postlethwait et al., 1998) and that their exact functions have not been well-studied. Although different at gene level, comparisons with the literature showed that in zebrafish and human HCC tumours similar families of genes and genes involved in shared pathways are altered in terms of methylation. For example, zebrafish HCC samples showed methylation changes in genes involved in proliferation, cell cycle, metastasis, apoptosis, energy production, adhesion, stress, DNA binding and regulation of transcription. Methylation pattern of genes involved in similar biological processes in human HCC were also altered (Tischoff and Tannapfel, 2008; Archer et al., 2010). Several other genes from families, such as ATP-binding cassette, sub-family A, DEAD-box helicases, potassium channel tetramerisation domain containing gene family and epidermal growth factor family were found to be hypermethylated in both zebrafish and human HCC (Gao et al., 2008). However, this work is preliminary. A more comprehensive study with same tumour causative agents in both species is required.

3.4.6. Conclusions

In this study development of a zebrafish CGI tiling microarray was described. It was demonstrated that in combination with the MeDIP technique, it can be used to detect and profile the methylation status of specific genes. This method was used to achieve a comprehensive profiling of DNA methylation in zebrafish liver as well as establishing the DNA methylation changes observed in HCC samples. Furthermore, the detected alterations in DNA methylation can help to explain some of the changes detected in gene expression in HCC samples. However, further functional studies are required to identify key CpG sites that are crucial in regulating gene expression. Nevertheless, striking similarities between the pathways disrupted by aberrant gene expression and DNA methylation were detected, both within and between human and zebrafish HCC.

Finally, achieving a better understanding of genetic and epigenetic regulation of transcription in zebrafish will increase the level of confidence in using zebrafish as a convenient model for investigating human diseases. Most importantly, this study has provided evidence of disruption of DNA methylation in critical pathways of tumourigenesis in fish.

Chapter 4

DNA methylation in liver tumourigenesis in dab (*Limanda limanda*) from the environment

The data presented in this chapter has been published.
Mirbahai et al., (2011) *Epigenetics*, 6 (11): 1319-1333.

4.1. Introduction

Traditionally cancer was considered as a multistep genetic disease driven and initiated by mutations (Vogelstein and Kinzler, 2004; Hanahan and Weinberg, 2000). However, data collected in the past decade demonstrate that both epigenetic and genetic changes interact and complement each other to enable the development of tumours (Gronbaek et al., 2007; Sharma et al., 2010; Feinberg, 2004).

The relationships between DNA methylation, chromatin modification and gene expression are complex. Several factors such as presence of transcriptional enhancers and suppressors, and their locations on the promoter, can subsequently influence the outcome (Siegfried and Simon, 2010; Jones 1999; see Chapter 1, section 1.3.2.3 and Chapter 3, section 3.4.4). In addition, the association between transcription and DNA methylation of the regulatory regions is further complicated by methylation of the coding regions of the genes. Genes that are moderately transcribed with no methylation in promoters have highly methylated coding regions compared to highly transcribed genes. This methylation has been linked to inhibition of transcription from cryptic sites located in the coding regions of the genes during transcription elongation (Suzuki and Bird, 2008; Zilberman et al., 2007). This suggests a higher susceptibility of the moderately transcribed genes to transcription interference (Zilberman et al., 2007). Nevertheless, irrespective of the specific influences of DNA methylation upon gene expression, distortion of the DNA methylation profile is a key event and a known hallmark of neoplastic cells with both global DNA hypomethylation and gene specific DNA hypermethylation (e.g. genes involved in apoptosis, metastasis, adhesion, cell cycle and DNA repair) reported in all cancers investigated (Lopez et al., 2009; Tischoff and Tannapfel, 2008; Gronbaek et al., 2007; Esteller and Herman, 2002; Esteller et al., 2001). Evidence is emerging that epigenetic changes, including alterations in DNA methylation, occur at early stages of tumourigenesis, potentially preceding mutational changes (Lopez et al., 2009; Pogribny, 2010; Reamon-Buthner et al., 2008; Feinberg et al., 2006; Sharma et al., 2010).

Environmental factors play a key role in the development of complex diseases. Exposures to natural and artificial chemicals and physical agents have been recognised as major contributors to the risk of development of cancers in humans (Parkin et al., 2002; Armstrong and Doll, 1975; for review see Wild, 2009). Previously, carcinogenic effects of chemicals

were mainly associated with the ability of the chemicals to interact with DNA and induce mutations (genotoxic chemicals). However non-genotoxic carcinogens have been identified as a second category that induce tumourigenesis but are not directly mutagenic. These chemicals modulate cell growth and proliferation through alteration of signalling pathways, including changes in DNA methylation; hence, resulting in changes to gene expression (Klaunig et al., 2000). Growing amount of experimental and epidemiological data demonstrate that accumulation of epigenetic modifications over time due to environmental insults (e.g. environmental stressors, including sub-lethal concentrations of chemicals), increases the individual's susceptibility to disease and development of disorders such as cancer (Reamon-Buthner et al., 2008; Wild, 2009; Bollati and Baccarelli, 2010; Baccarelli and Bollati, 2009; Feinberg et al., 2006). This correlates with the considerable time required for development of different types of tumours (Baccarelli and Bollati, 2009; Feinberg et al., 2006). For example, both rodent and epidemiological studies on exposure to metals (e.g. nickel, cadmium, chromium, arsenic), air pollutants and endocrine disruptors have clearly demonstrated a link between environment, epigenome and cancer (for reviews see Skinner et al., 2010; Arita and Costa, 2009; Baccarelli and Bollati, 2009; Chapter 1, section 1.3.5.2).

However, in contrast to the wealth of information available on environmental factors, change in the epigenome and development of cancer in humans, this area is substantially understudied in aquatic biology. A few studies have investigated the effects of chemical exposures on DNA methylation levels at either global level or on selected genes in aquatic species. These are; hypomethylation of the vitellogenin 1 promoter in adult zebrafish liver (*Danio rerio*) following exposure to 17 α -ethinylestradiol (EE) (Stromqvist et al., 2010), global DNA hypermethylation in male gonads of three-spined stickleback (*Gasterosteus aculeatus*) after exposure to 17 β -estradiol (E2) (Aniagu et al., 2008), global DNA hypermethylation in the liver of goldfish (*Carassius auratus*) following exposure to heavy metals (Zhou et al., 2001), changes in methylation of aromatase and estrogen receptor (ER) promoters in Japanese medaka (*Oryzias latipes*) following exposure to EE (Contractor et al., 2004), genome-wide methylation profiling of adult zebrafish hepatocellular carcinoma tumours induced by DMBA (our own study, Mirbahai et al., 2011, Chapter 3), global DNA methylation studies in *Daphnia magna* following exposure to several chemicals (Vandeghechuchte et al., 2009a, Vandeghechuchte et al., 2009b, Vandeghechuchte et al., 2010a; Vandeghechuchte et al., 2010b) and global DNA methylation studies in the liver of bluegill

sunfish (*Lepomis macrochirus*) and kelpfish (*Sebastiscus marmoratus*) following exposure to benzo[*a*]pyrene, tributyltin and triphenyltin (Shugart et al., 1990; Wang et al., 2009) (for more details on each study see Chapter 1, section 1.3.1, Table 1.1). As a result, it is becoming more apparent that epigenetic mechanisms in both aquatic species and mammals are affected by model environmental pollutants. This highlights the importance of investigating the influence of environmental factors upon the epigenome and determining its role in diseases of marine species. However, so far there have been no studies of marine species investigating epigenetic changes in relation to carcinogenesis, partly due to the difficulties of studying non-model organisms and the availability of tumour samples from wild species.

Common dab (*Limanda limanda*) is a flatfish caught from UK waters as part of the UK Clean Seas Environmental Monitoring Programme (CSEMP). Liver pathology in dab, including cancer and pre-neoplastic lesions, is used as an indicator of the biological effects of contaminants on the marine environment. Due to their living habits and close proximity to the ocean floor they can be exposed to relatively high levels of sediment-associated chemicals. This makes them ideal species for biomonitoring and environmental carcinogenicity research. An unusually high prevalence of liver tumours has been reported in dab, with some UK sites exceeding 20% of sampled individuals (NMMP, second report, 2004; Stentiford et al., 2009; Small et al., 2010; Southam et al., 2008; Lyons et al., 2006). However, the causative factors of these tumours and the molecular mechanisms involved, especially in relation to the balance between epigenetic and genetic factors, is unclear. Previous studies in our laboratory indicated that mutation profiles of oncogenes and tumour-suppressor genes within tumours were markedly different in frequency between fish species and humans (Rotchell et al., 2001a; Franklin et al., 2000). In Chapter 3 and in our publication (Mirbahai et al., 2011) we demonstrated that pathways with differentially methylated genes in chemically induced zebrafish (*D. rerio*) hepatocellular carcinoma (HCC) are similar to the pathways altered in human HCC. This highlighted the importance of DNA methylation alterations in the development of fish tumours and suggested the involvement of epigenetic factors in the formation of dab tumours.

Therefore, having confirmed the alteration of DNA methylation in fish tumours, based on the results of Chapter 3, the aims of this study were to identify differentially methylated regions of the genome as well as differentially expressed genes in dab hepatocellular adenoma (HCA)

compared to healthy dab liver. In addition, using pathway analysis techniques we aimed to determine if methylation and expression of genes in specific pathways were altered in tumours compared to healthy control samples. The ultimate goal was to elucidate possible molecular mechanisms behind these tumours and determine if a link could be established between environmental contaminants, epigenetic changes and dab liver tumours.

Dab is an un-sequenced species, therefore using traditional techniques for gene-specific studies in dab is challenging. Hence, to overcome this problem and to identify genome-wide methylation changes we combined MeDIP with *de novo* high-throughput DNA sequencing for the first time. Based on this data, we designed the first dab-specific gene expression microarray. Finally, this data demonstrates the suitability and power of these methods for studying DNA methylation in un-sequenced species.

4.2. Overview of the experimental approach (Methodological details are described in Chapter 2)

The HPLC method, established in Chapter 3, was used for measurement of global DNA methylation levels in three different categories of interest, dab hepatocellular adenoma (HCA), healthy tissue surrounding HCA samples (ST) and dab healthy liver. Samples (healthy liver, ST and HCA) used for this study were all confirmed by histopathology studies.

MeDIP combined first with a cDNA microarray as a pilot experiment, then with high-throughput sequencing (HTS) were used to investigate alterations in DNA methylation in HCA samples compared to non-cancerous surrounding tissue (MeDIP-HTS was carried out by Beijing Genomics Institute (BGI)). In addition, bisulfite sequencing PCR was used to confirm the observed changes in DNA methylation.

A dab-specific 8x15k gene expression microarray was constructed based on the sequences achieved from the HTS of a HCA sample and corresponding ST. This microarray was used for transcriptomic profiling of the three sample categories of interest. Real time-PCR was used for validation of the data achieved from the transcriptomic study.

GeneSpring, multiexperiment viewer (MeV), Blast2GO, Ingenuity Pathway Analysis and PCA were used to analyse and identify the biological functions that were altered in HCA samples compared to ST and healthy dab liver.

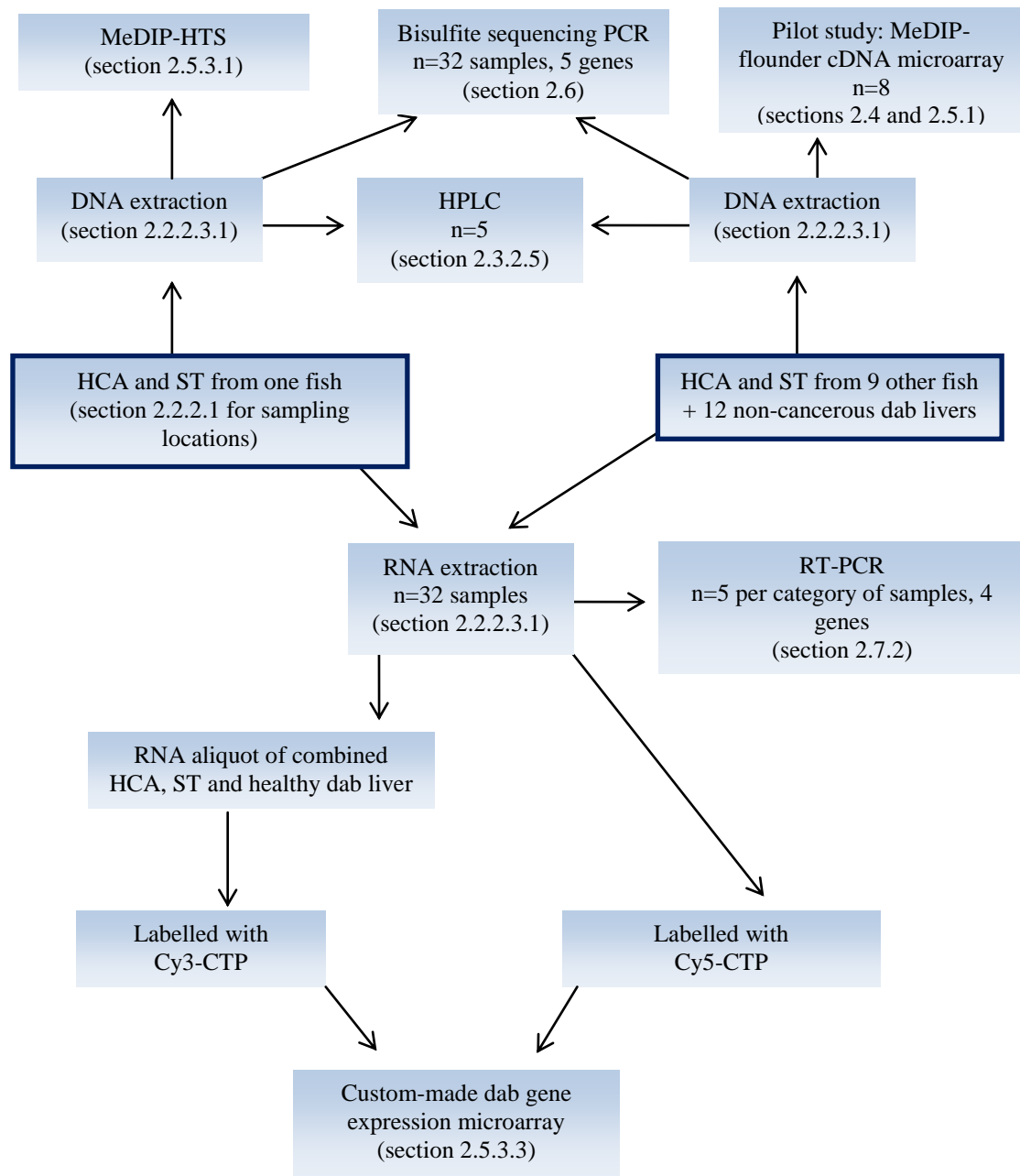


Figure 4. 1. Flowchart of the experimental procedures used in Chapter 4. The starting points of the diagram are the boxes with borders. The methods used in this chapter are described in detail in Chapter 2. In this flowchart, the sections where these methods are described in Chapter 2 are shown in brackets.

4.3. Results

4.3.1. Global measurement of genomic DNA methylation

To establish the global levels of cytosine methylation in the three tissue categories of interest (healthy dab liver, HCA and apparently healthy surrounding tissue of tumour containing dab liver (ST)) HPLC was performed (number of replicates for each category=5). Absence of RNA contamination was confirmed as previously described (Chapter 3, section 3.3.1.2). As shown in Figure 4.2 a statistically significant 1.8-fold DNA hypomethylation was detected between healthy dab liver and tissue surrounding HCA tumours (ST) and between healthy dab liver and HCA ($P<0.01$). However, no significant differences were detected in overall cytosine methylation levels between ST and HCA. The level of cytosine methylation in healthy dab liver (Figure 4.2) was 8%. This is in agreement with previous published data demonstrating a 2-fold higher DNA methylation level in fish liver compared to the liver of mammals (Aniagu et al., 2008).

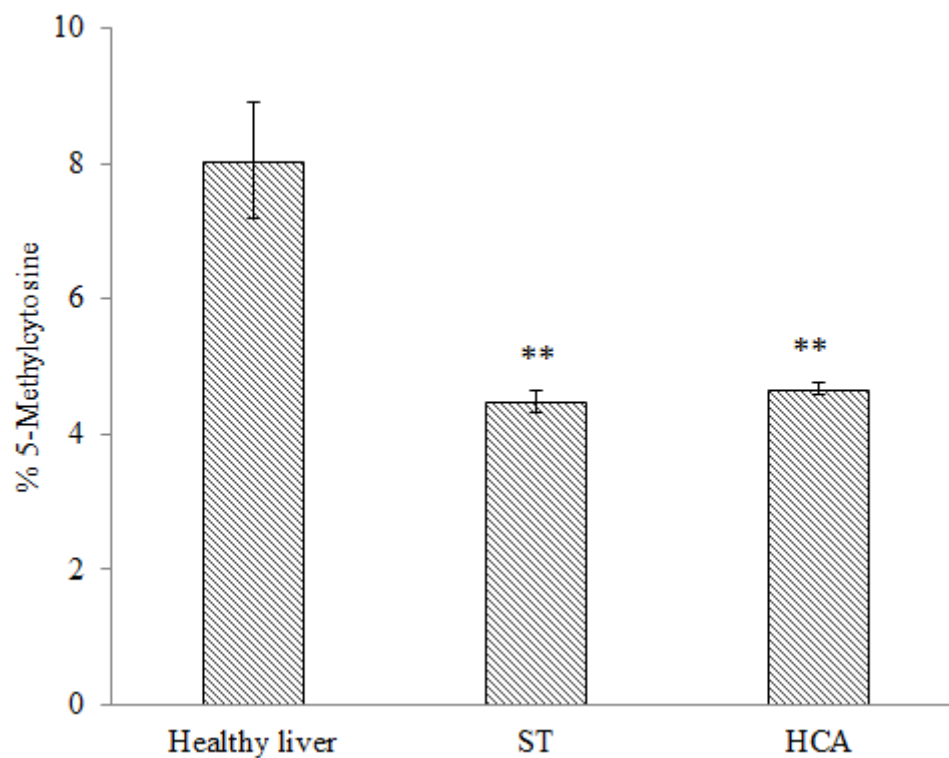


Figure 4.2. Measurement of global percentage of methylated cytosine. Overall DNA methylation levels were measured in healthy dab liver, hepatocellular adenoma (HCA) and tissue surrounding HCA tumours in dab liver (ST) using HPLC. Overall methylation levels of HCA and ST differed significantly from healthy liver. No significant difference was detected in overall DNA methylation levels between HCA and ST. This was investigated using one-way ANOVA Tukey's post-hoc test. ** Significantly different from healthy liver ($P < 0.01$). Data are shown as mean \pm SEM of five independent samples.

4.3.2. DNA methylation at gene levels in dab liver

4.3.2.1. Pilot study using MeDIP coupled with flounder cDNA microarray

Although no differences were detected in overall cytosine methylation levels between ST and HCA, to investigate further the possibility of differentially methylated regions between HCA and ST at the gene level, MeDIP was combined with a flatfish European flounder (*Platichthys flesus*) cDNA microarray as described previously (Chapter 2, section 2.5.1). DNA samples were extracted from 8 HCA and corresponding surrounding tissue (ST) of female dab and sonicated (sample details are described in Chapter 2, section 2.2.2.1) then immunoprecipitated as described previously (Chapter 2, section 2.4) and labelled with Cy5-dCTP. These samples were hybridised to a flounder cDNA microarray against a reference sample comprised of sonicated genomic DNA pool labelled with Cy3-dCTP.

The flounder cDNA microarray was constructed as part of the Genomic Tools for Biomonitoring of Pollutant Coastal Impact (GENIPOL) project in 2006 (Williams et al., 2006) containing 12,738 clones spotted in duplicate (either expressed sequenced tags (ESTs) or cloned regions of genes) relating to approximately 3336 unique coding sequences from the European flounder. Sequence homology is sufficiently great between the two closely related flatfish; therefore, the use of this microarray for transcriptional studies in dab is valid (Diab et al, 2008; Cohen et al., 2007; Small et al., 2010).

Two fish samples were excluded from the analysis. One HCA and matching ST sample were removed from the analysis as the microarray slide containing the HCA sample (sample code: RA08045-43) broke prior to scanning of the slide. The second set was removed during sample quality checks (sample code: RA08045-10). The HCA sample showed poor hybridisation to the microarray probes and therefore along with the matching ST was excluded during data analysis. This reduced the sample numbers to 6 pairs of HCA and ST.

Following quality checks and normalisation, the data were analysed using Welch paired *t*-test with assumption of un-equal variance ($P < 0.05$) as described previously (Chapter 2, section 2.9.2). This approach resulted in identification of 78 clones (15 hypomethylated and 63 hypermethylated) that were statistically significantly altered in HCA samples compared to the non-cancerous ST with 16 clones (3 hypomethylated and 13 hypermethylated) showing

altered methylation levels of greater than 1.5-fold (Table 4.1; complete list of all potentially differentially methylated clones are presented in Appendix File 7).

Due to several reasons, explained in the discussion, this study resulted in identification of small numbers of differentially methylated clones in HCA compared to ST (>1.5-fold change). The data achieved from this pilot study were only used as a qualitative approach for investigating the possibility of alteration in DNA methylation at gene level between HCA and ST rather than a comprehensive quantitative study.

4.3.2.1.1. Principal components analysis

From the two categories of interest, all the genes that differed with disease status were subjected to PCA scores plot (Figure 4.3). HCA and ST samples were separated based on treatment with PC1 and PC2 accounting for 84.25% and 6.25% of the variations observed, respectively.

Function	Common name	Fold change	P-value
Antioxidant activity	<i>glutathione peroxidase 1</i>	-1.59	0.03
Antioxidant activity	<i>peroxiredoxin 6</i>	-1.43	0.03
Protein synthesis	<i>60S acidic ribosomal protein P0 (L10E)</i>	1.28	0.045
Protein synthesis	<i>mitochondrial ribosomal protein S18C</i>	1.37	0.019
Protein synthesis	<i>ribosomal protein L19</i>	1.27	0.039
Drug transporter	<i>multidrug resistance protein (mdr gene)</i>	1.38	0.045
Energy and metabolism	<i>aldolase B</i>	1.40	0.041
Amino Sugar metabolism	<i>N-acylglucosamine 2-epimerase</i>	1.41	0.024
Anti-proliferation activity	<i>retinoic acid receptor responder</i>	1.60	0.006
Viral transcription	<i>reverse transcriptase-like protein</i>	1.52	0.006
Virus related gene (fish disease)	<i>viral hemorrhagic septicemia-induced protein-5</i>	1.34	0.037
Angiogenesis	<i>vascular endothelial growth factor</i>	1.32	0.017
Migration and growth	<i>ras homolog gene family, member A</i>	1.48	0.016
Protein degradation	<i>proteasome (prosome, macropain) 26S subunit , 12</i>	-1.28	0.002
Protein degradation	<i>chymotrypsinogen 1</i>	-1.15	0.015

Continued from previous page

Lipid metabolism	<i>apolipoprotein AI</i>	-1.28	0.029
Structural role	<i>protein transport protein Sec61 alpha subunit</i>	-1.77	0.04
Structural role	<i>actin</i>	1.30	0.046
Structural role	<i>collagen, type XII, alpha 1</i>	1.72	0.021
Apoptosis	<i>B-cell receptor-associated protein 31</i>	1.39	0.045
Immune system, inflammation	<i>myeloid/lymphoid or mixed-lineage leukaemia protein 4</i>	1.37	0.012
Immune system, inflammation	<i>CD3G antigen, gamma polypeptide (TIT3 complex)</i>	1.33	0.013
Immune system, inflammation	<i>T-cell leukaemia translocation altered gene (TCTA)</i>	1.27	0.039
Immune system, stem cell regeneration, inflammation	<i>mesenchymal stem cell protein DSCD75</i>	1.56	0.006

Table 4.1. Subsection of the identified clones with potentially altered methylation levels in dab HCA samples compared to ST determined by MeDIP-cDNA flounder microarray. The biological function, common name, fold change (tumour versus ST) and the *P*-value corresponding to each clone are shown in this table.

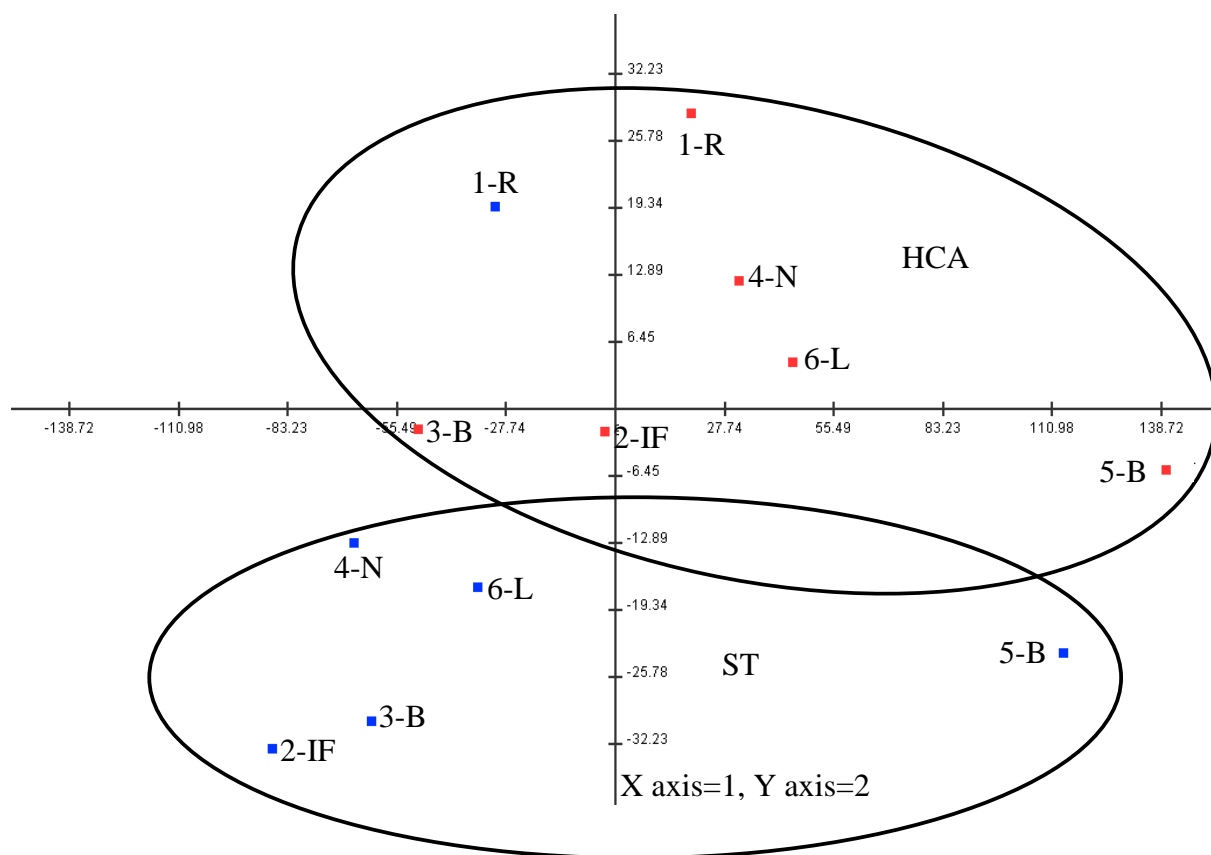


Figure 4.3. Principal component analysis (PCA) scores plot for all the genes that were differentially methylated in the HCA samples compared to ST in MeDIP-cDNA microarray data. Dab HCA samples (red) and surrounding tissue samples (blue) were separated based on treatment along the PC1 and PC2 axes. One ST sample is grouped with the HCA samples. Numbers 1-6 represent 6 individual fish. Matching numbers are ST and HCA samples from the same fish. Sampling sites for each fish is also shown. N: North Cardigan Bay, B: St Bee's Head, R; Red Wharf Bay, L: Lundy, IF: Indefatigable Bank.

4.3.2.2. Comprehensive DNA methylation profiling of dab HCA tumours

4.3.2.2.1. *De novo* high-throughput sequencing analysis of MeDIPed DNA

The pilot study using MeDIP coupled with a flounder cDNA microarray provided a reasonable level of confidence that DNA methylation at gene level varied between HCA and ST samples collected from the livers of flatfish dab. Therefore to achieve a comprehensive, dab sequence-specific study and to overcome the complications encountered in the pilot study, (see discussion), *de novo* high-throughput sequencing (HTS) was combined with MeDIP, for the first time, as described in Chapter 2, section 2.5.3.1. The genome-wide DNA methylation profiling of HCA and corresponding ST was performed by first generating MeDIP enriched DNA libraries. Immunoprecipitated fragments were subjected to high-throughput sequencing using an Illumina Genome Analyser II to obtain a comprehensive DNA methylation map (Raw data for both samples, hepatocellular adenoma HCA-18-RA09065-830 and healthy surrounding tissue ST-19-RA09065-830, were submitted to NCBI database and can be found under accession numbers GSE31124 and GSM770685, respectively). The 89,925,735 and 87,826,470 44-base paired-end reads for HCA and ST respectively, were combined and assembled. This resulted in identification of 264,008 contigs with lengths greater than or equal to 200 bp. From the identified contigs, the ones that did not match the criteria for prediction of CGIs in vertebrates, (criteria for prediction of CGI: length ≥ 200 bp, O/E ≥ 0.6 , CG% ≥ 50) as described by Gardiner-Garden and Frommer (Gardiner-Garden and Frommer, 1987), were excluded (Table 4.2).

Contigs criteria	Total number	Total length	Average length	Maximum length
Length ≥ 200 bp	264,008	78,492,996	297	2,283
Length ≥ 200 bp, O/E ≥ 0.6 , CG% ≥ 50	69,046	22,317,135	323	2,283

Table 4.2. CpG island discovery. The reads obtained from MeDIP high-throughput sequencing of HCA and corresponding surrounding tissue were combined. The criteria described by Gardiner-Garden and Frommer (1987) were used to obtain 69,046 CGI-containing contigs for further analysis and identification of differentially methylated regions between HCA and surrounding tissue.

In order to classify each segment as hyper-, hypo-, or non-differentially methylated, the number of MeDIP-reads comprising each contig in tumour were divided by the number of MeDIP-reads that comprised the same contig in corresponding ST. Hypo- and hyper-methylated regions in HCA compared to ST were defined arbitrarily as 1.5-fold or greater decrease and increase in the calculated ratio, respectively. As dab is an un-sequenced species, the identified candidate hypo- and hyper-methylated contigs were annotated against the Genbank non-redundant (nr) protein database (E-value cut-off $<10^{-6}$). This resulted in identification of 1,693 differentially methylated sequences containing CGI regions in tumour compared to control (Figure 4.4; a complete list of identified contigs (1,693), annotations and fold changes are shown in Appendix File 8). To identify potentially differentially methylated CGI-containing promoter regions in dab HCA compared to ST, the identified contigs (69,046) were mapped onto pufferfish (*Takifugu rubripes*) genome available from UCSC genome browser website. Pufferfish was selected for contig mapping due to its higher sequence homology to dab compared with other fully sequenced fish species (zebrafish, stickleback, medaka and Atlantic salmon (*Salmo salar*)). This was decided based on the initial annotation that was conducted against the Genbank nr protein database. This resulted in identification of 60 hypo- and 51 hyper-methylated (>1.5 -fold change, Figure 4.4) CGIs that could be aligned with the pufferfish genome between 1.5 kb upstream and 1 kb downstream of the predicted transcription start sites. One of the key genes identified in this list was *DNA (cytosine-5)-methyltransferase 3 beta* which was hypomethylated in HCA compared to ST.

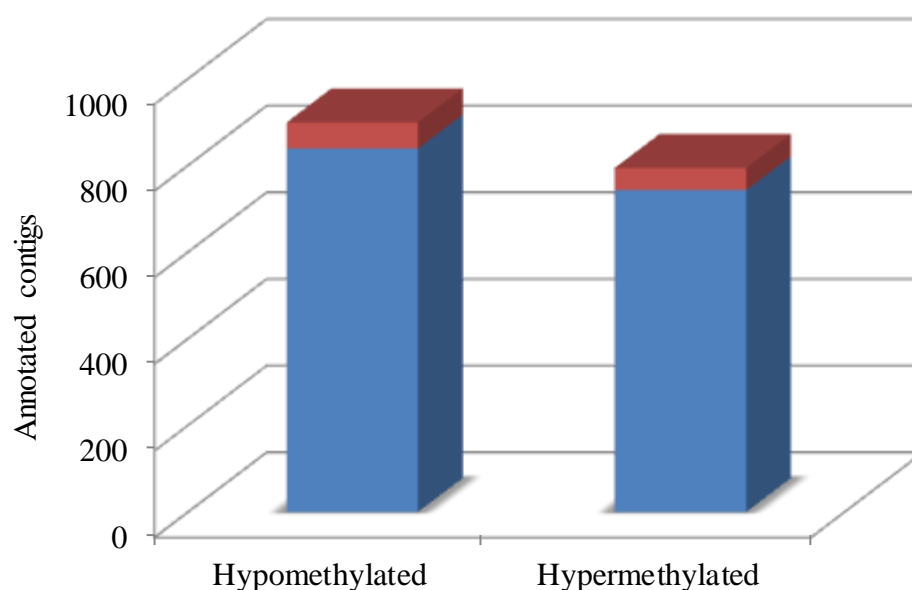


Figure 4.4. Identification of differentially methylated contigs in dab HCA compared to surrounding tissue. Combined blue and red sections: 899 hypomethylated and 794 hypermethylated CGIs in dab HCA compared to ST that could be annotated against Genbank nr protein database (>1.5 -fold change, E-value cut-off $<10^{-6}$). Red sections: 60 hypomethylated and 51 hypermethylated CGIs identified from dab that aligned to 1.5 kb upstream to 1 kb downstream of the predicted transcription start sites (TSS) in pufferfish genome (>1.5 -fold change; the list of 111 hypo- and hyper-methylated regions in TSS are shown in Appendix, File 9).

4.3.2.2.2. Ingenuity Pathway Analysis

Ingenuity Pathway Analysis (IPA) was performed to characterise the functional relationships between genes with altered methylation in HCA compared to ST. The most significant networks and functions included cell-to-cell signalling, cell cycle, DNA replication and cellular assembly and organisation as well as canonical pathways including Wnt/ β -catenin signalling, growth hormone signalling and apoptosis signalling. Figure 4.5 represents the IPA biological functions associated with genes with altered methylation levels in HCA (FDR $<5\%$). Table 4.3 illustrates a sub-section of the genes (hypo- and hyper-methylated) possessing biological functions linked to carcinogenesis.

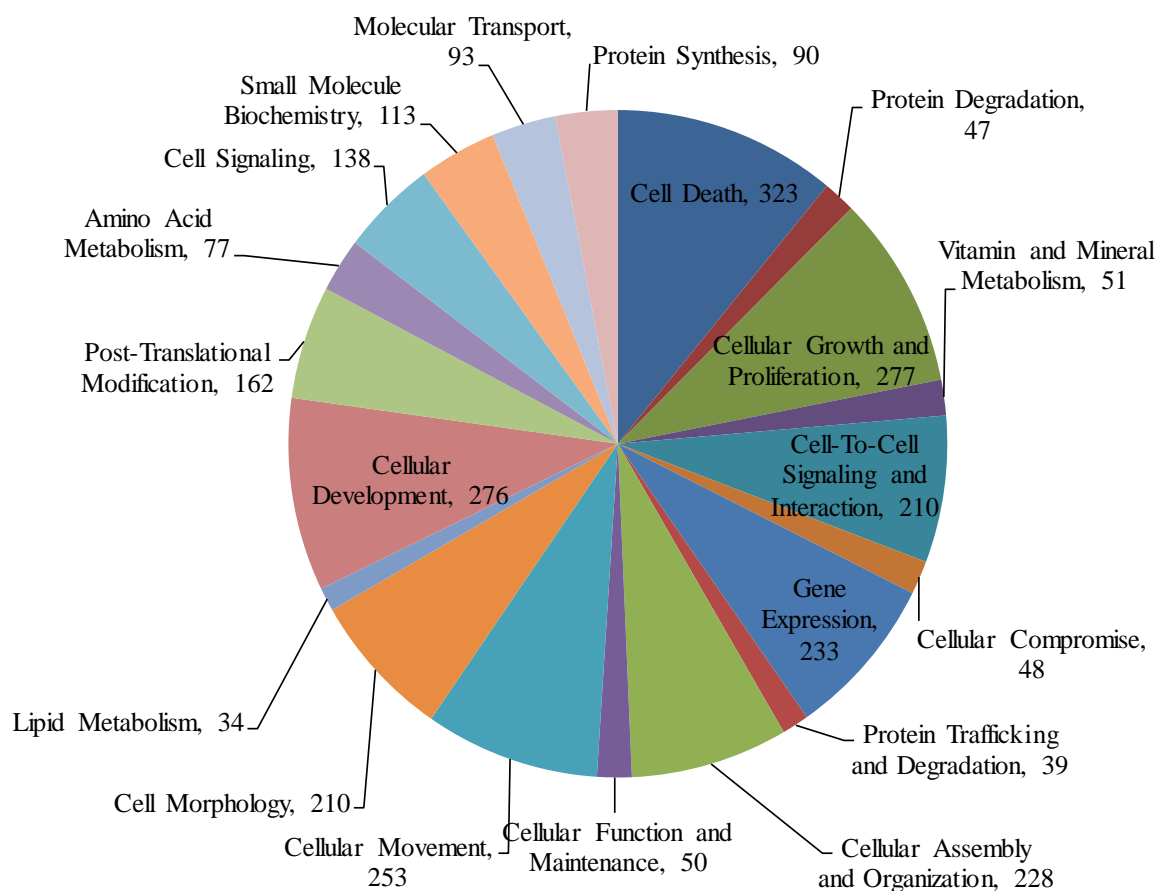


Figure 4.5. Biological functions enriched among genes with altered methylation levels (hypo- and hyper-methylated) in HCA. IPA was used to group the genes with altered methylation levels based on biological functions (>1.5 -fold change, $FDR < 5\%$). The numbers of genes associated with each biological function are shown.

Biological function	Methylation level in HCA compared to ST
Cell death (cell death, apoptosis, survival of cells)	<i>caspase 6 (casp6,↑), cell death-inducing dffa-like effector b (cideb,↓), dedicator of cytokinesis 1 (dock1,↑), apoptosis-associated tyrosine kinase (aatk,↓), angiopoietin-related protein 4 precursor (angptl4,↓), synovial apoptosis inhibitor synoviolin (syvn,↑), signal transducer and activator of transcription 5 (stat5a,↓), jagged 1 (jag1,↓), jagged 2 (jag2,↓), rho-associated protein kinase 1 (rock1,↑), cyclin-dependent kinase 6 (cdk6,↑)</i>
Cell growth and proliferation	<i>fibroblast growth factor receptor 3 (fgfr3,↓), insulin-like growth factor 1 receptor (igf1r,↓), jagged 1 (jag1,↓), jagged 2 (jag2,↓), transcription factor e2f7 (e2f7,↓), DNA (cytosine-5-)-methyltransferase 3b (dnmt3b,↓), signal transducer and activator of transcription 5 (stat5a,↓), glucose-6-phosphate 1-dehydrogenase (g6pd,↓), discs, large homolog 1 (dlg1,↓), peroxisome proliferator-activated receptor beta (pparb,↓)</i>
Cell-to-cell signalling and interaction	<i>jagged 1 (jag1,↓), jagged 2 (jag2,↓), signal transducer and activator of transcription 5 (stat5a,↓), cadherin 22 (cdh22,↓), integrin, alpha 3a (itga3,↑), integrin beta 1 subunit (itgb1,↑), integrin alpha 9 (itga9,↑), protocadherin 1 gamma 22 (pcdh1g22,↓)</i>
Gene expression	<i>cAMP responsive element binding protein 1 (crem,↑), notch homolog 2 (notch2,↓), aryl hydrocarbon</i>

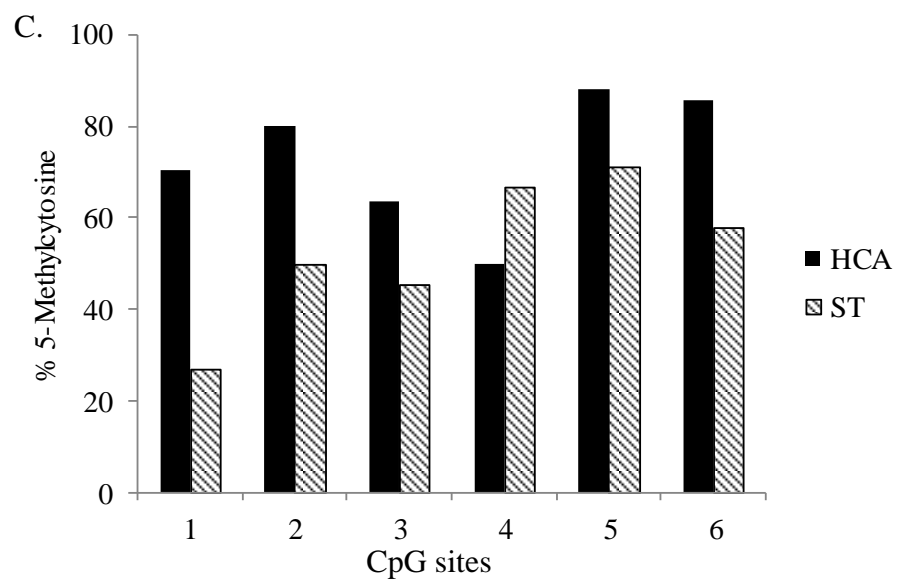
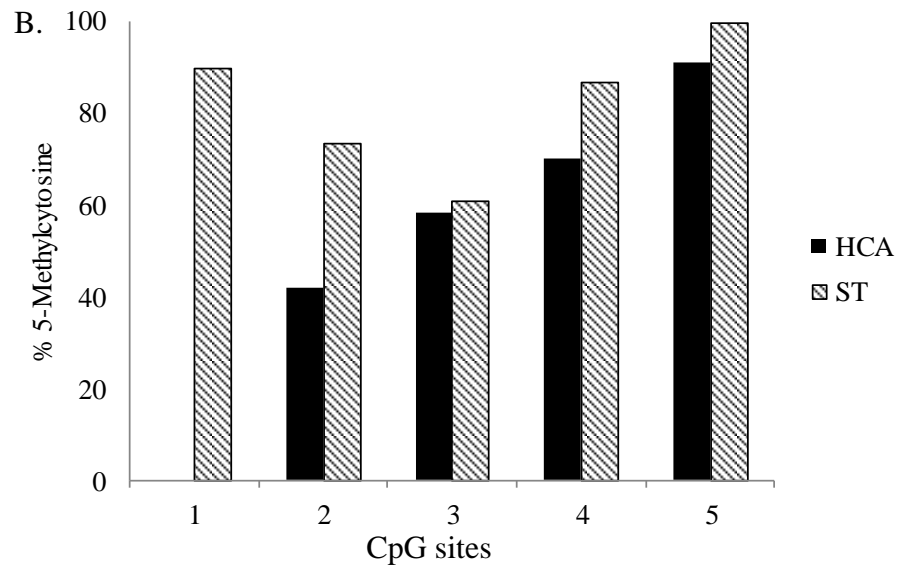
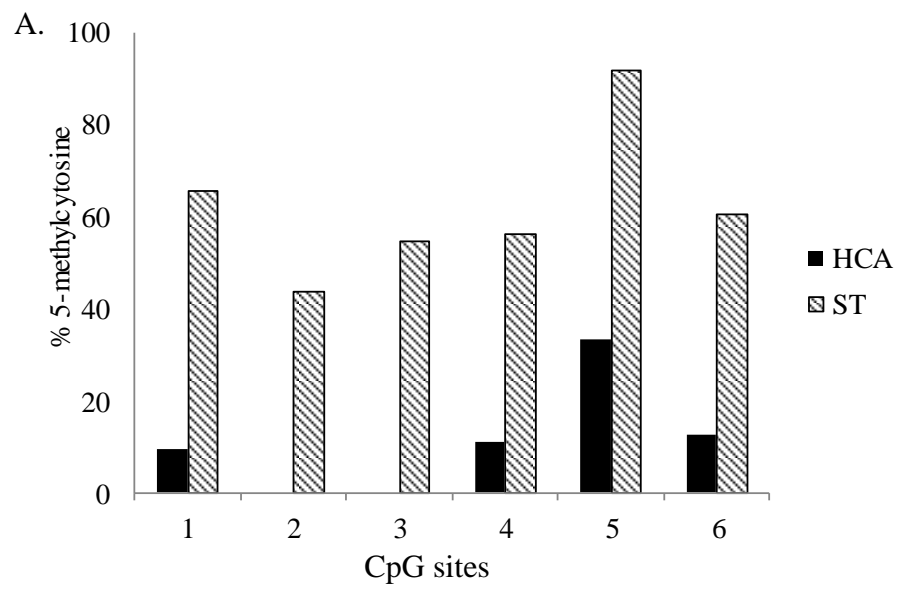
Continued from previous page

Gene expression	<i>receptor nuclear translocator-like protein 1(arntl,↑), hematopoietic transcription factor gata-1 (gata1,↑), mitogen-activated protein kinase kinase kinase kinase 4-like isoform 3 (mapk4k3,↑), mitogen-activated protein kinase 12 (mapk12,↓), mitogen-activated protein kinase 8 interacting protein 3 (mapk8ip3,↑), mitogen-activated protein kinase 11 (mapk11,↓), mitogen-activated protein kinase 1 (mapk1,↑)</i>
-----------------	--

Table 4.3. Subsection of the genes with altered methylation associated with development of tumours (FDR <5%) in HCA compared to ST. ↑ hypermethylated and ↓ hypomethylated in HCA compared to ST. The complete list of the biological functions and associated genes are presented in Appendix File 10.

4.3.2.2.3. Confirmation of the MeDIP *de novo* high-throughput sequencing data using BSP and comparability of the data to additional individuals

To validate the data achieved from MeDIP *de novo* HTS, five differentially methylated genes were selected with a fold change of greater than 2-fold, including both hypomethylated and hypermethylated groups of genes and an unidentified sequence. Direct bisulfite sequencing PCR (see Chapter 3, section 3.3.3) was carried out on the original samples used for MeDIP *de novo* HTS, as previously described. Cytosines located at CpH (H=A,T,C) sites were used to monitor the efficiency of bisulfite treatment and sequencing. Efficient bisulfite treatment should result in 100% conversion of the cytosine base to T at these regions. The data for the 5 genes investigated, *protocadherin 1 gamma 22* (*pcdh1g22*), *microtubule-associated protein 1aa* (*map1aa*), *tubulin tyrosine ligase-like member 7* (*tll7*), *nidogen 1* (*nid1*) and an unidentified sequence, are shown in Figure 4.6. For all the genes investigated the BSP data were in complete agreement with the data from MeDIP high-throughput sequencing.



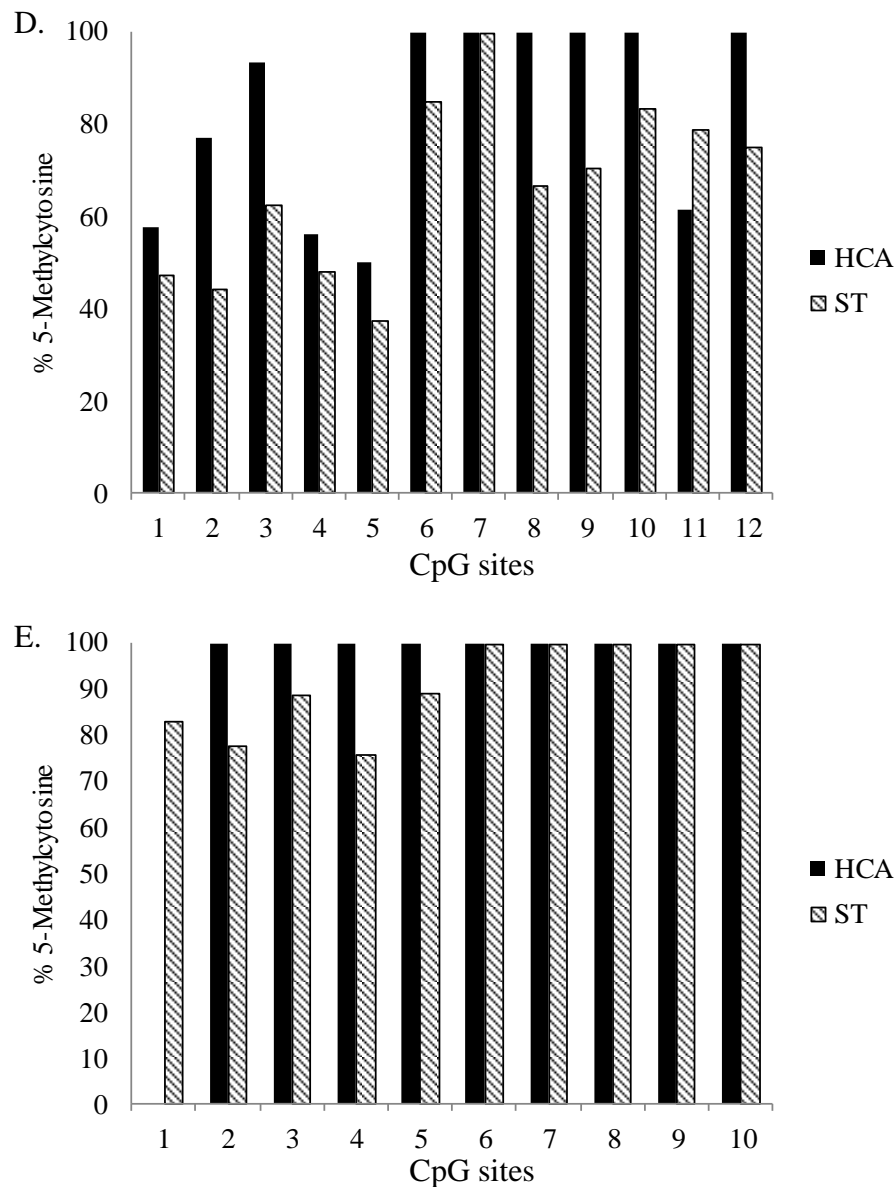


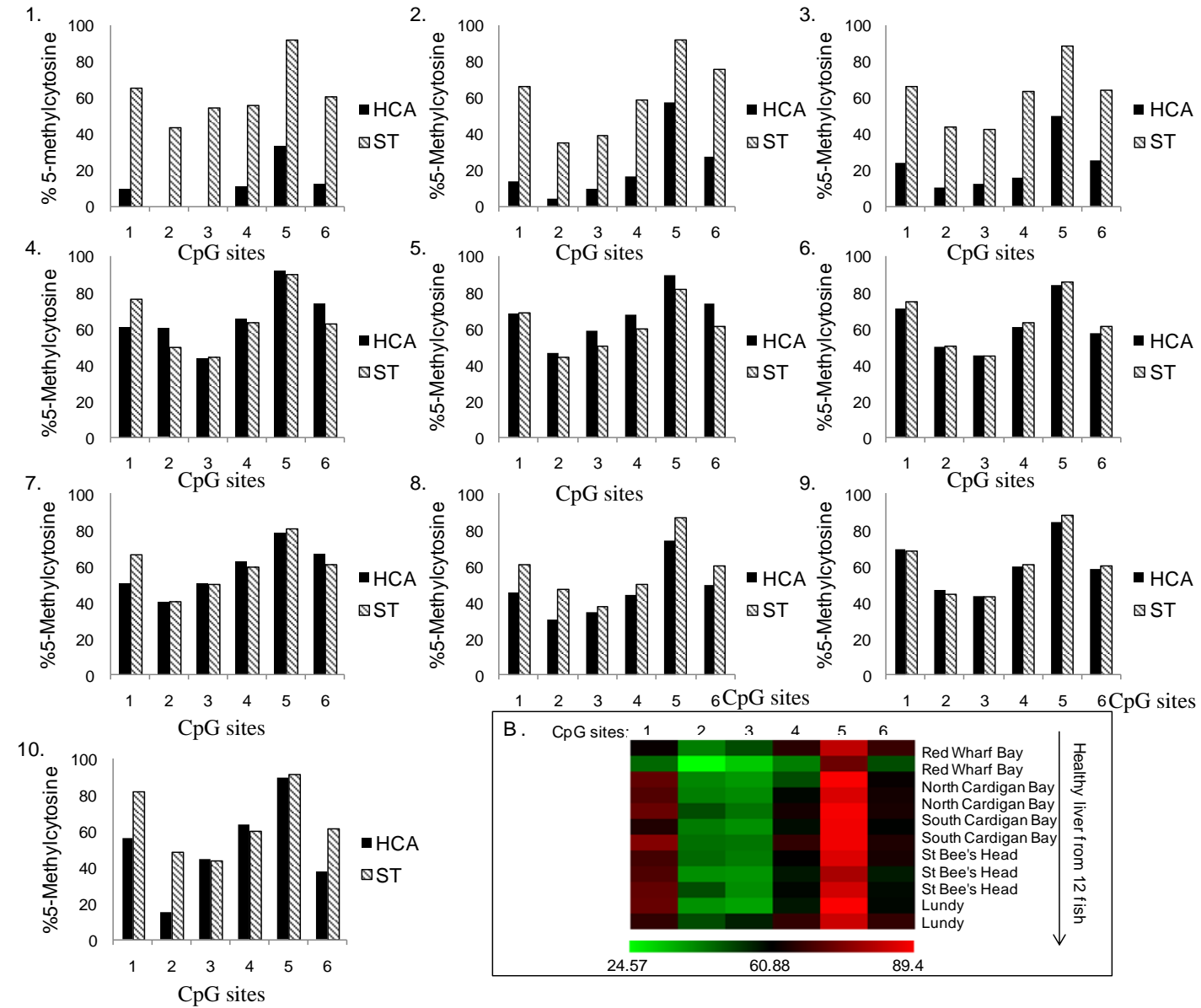
Figure 4.6. Direct bisulfite sequencing PCR data for A. *protocadherin 1 gamma 22* (*pcdh1g22*, -2.08), B. *microtubule-associated protein 1aa* (*map1aa*, -7.69), C. an unidentified sequence (+4.17), D. *tubulin tyrosine ligase-like member 7* (*ttl7*, +3.05) and E. *nidogen 1* (*nid1*, +2.01). To validate the MeDIP high-throughput sequencing data the same HCA and corresponding ST were used for direct BSP. The numbers in front of the gene symbols in the brackets demonstrate the fold change in DNA methylation level detected by MeDIP HTS for the mentioned gene for the same region investigated with direct BSP. +: hypermethylation, -: hypomethylation, numbers on the x-axis: represent CpG sites. (x axis label added). Absence of a bar indicates unmethylated CpG sites.

4.3.2.2.4. Comparison of the BSP data

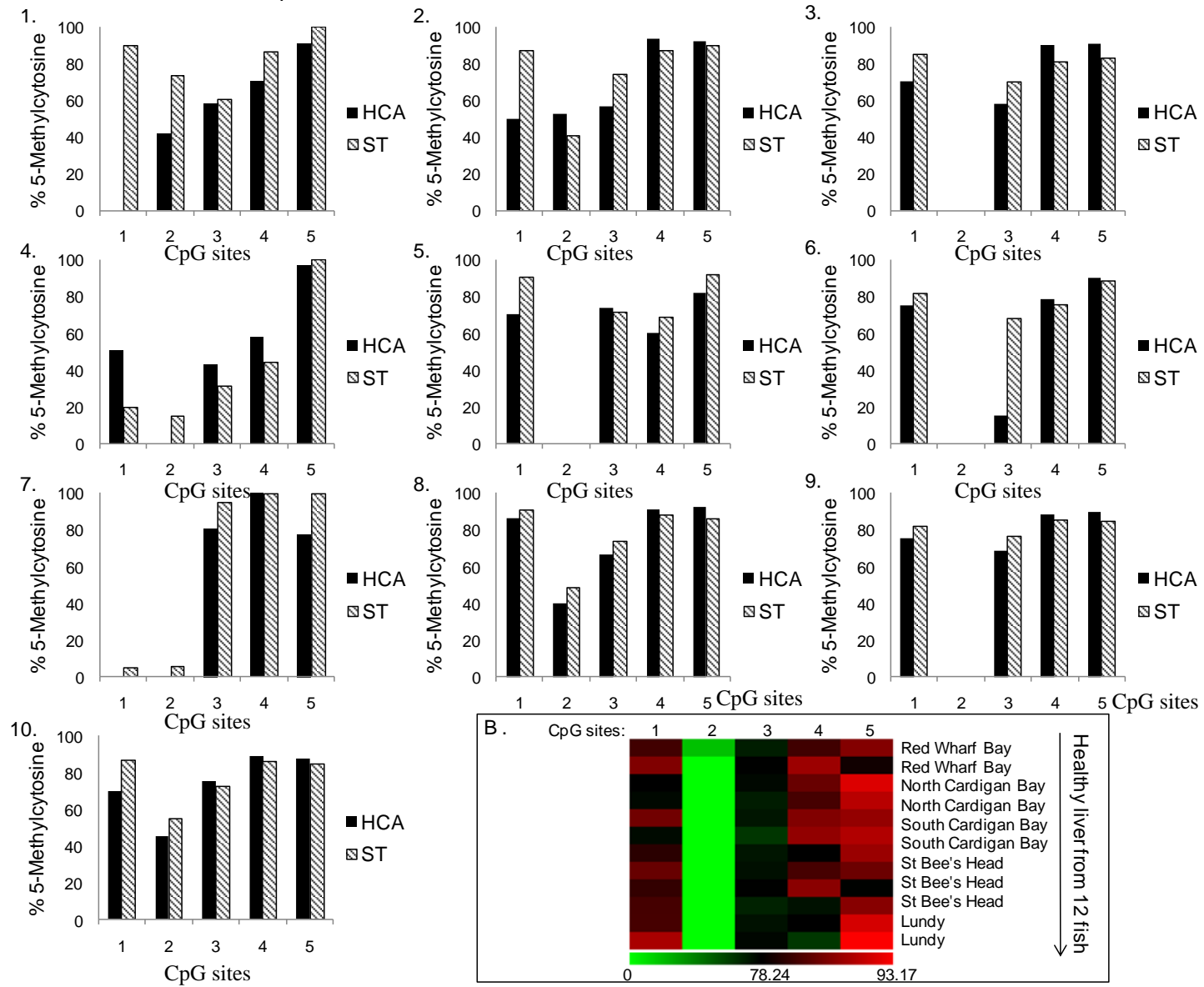
In addition, the levels of variation and comparability of the DNA methylation data achieved from MeDIP *de novo* HTS and BSP of the one fish described in section 4.3.2.2.3, to HCA and corresponding ST from 10 individual dab collected from five different sampling sites in the Irish Sea and Bristol Channel were investigated using BSP (sampling details are shown in Chapter 2, section 2.2.2.1). To achieve this, BSP was conducted for the same five genes mentioned in section 4.3.2.2.3 on 10 HCA, 10 corresponding ST and 12 healthy tissues (Figure 4.7). To clarify the figures, the data for healthy fish liver samples are presented as heat maps. However, as the main focus is on tumour and ST samples the data for these samples are presented as individual graphs.

The same trends in methylation levels were observed between HCA and surrounding tissue samples for most genes and fish investigated. However, inter-individual variation was observed, especially between HCA samples of different fish. The lowest level of variation in DNA methylation profile was observed between different healthy dab livers. The observed variation was not surprising as these were wild fish, exposed to different environmental conditions, such as different levels of contaminants. Furthermore, from the pathology data for these samples it is evident that these tumours are heterogeneous in their stage of development and other underlying conditions (pathology data are presented in Appendix File 11). For example, different HCA samples demonstrated various stages of hepatic hydropic vacuolation (HydVac). This condition is characterised by cells with large non-staining vacuoles and small peripheral nuclei. In certain fish species including flatfish, such as winter flounder (*Pseudopleuronectes americanus*) this condition has been associated with exposure to xenobiotics (e.g. polychlorinated biphenyls (PCBs) and polycyclic aromatic hydrocarbons (PAHs)) and neoplasia (Moore et al., 1997; Stehr et al., 1998).

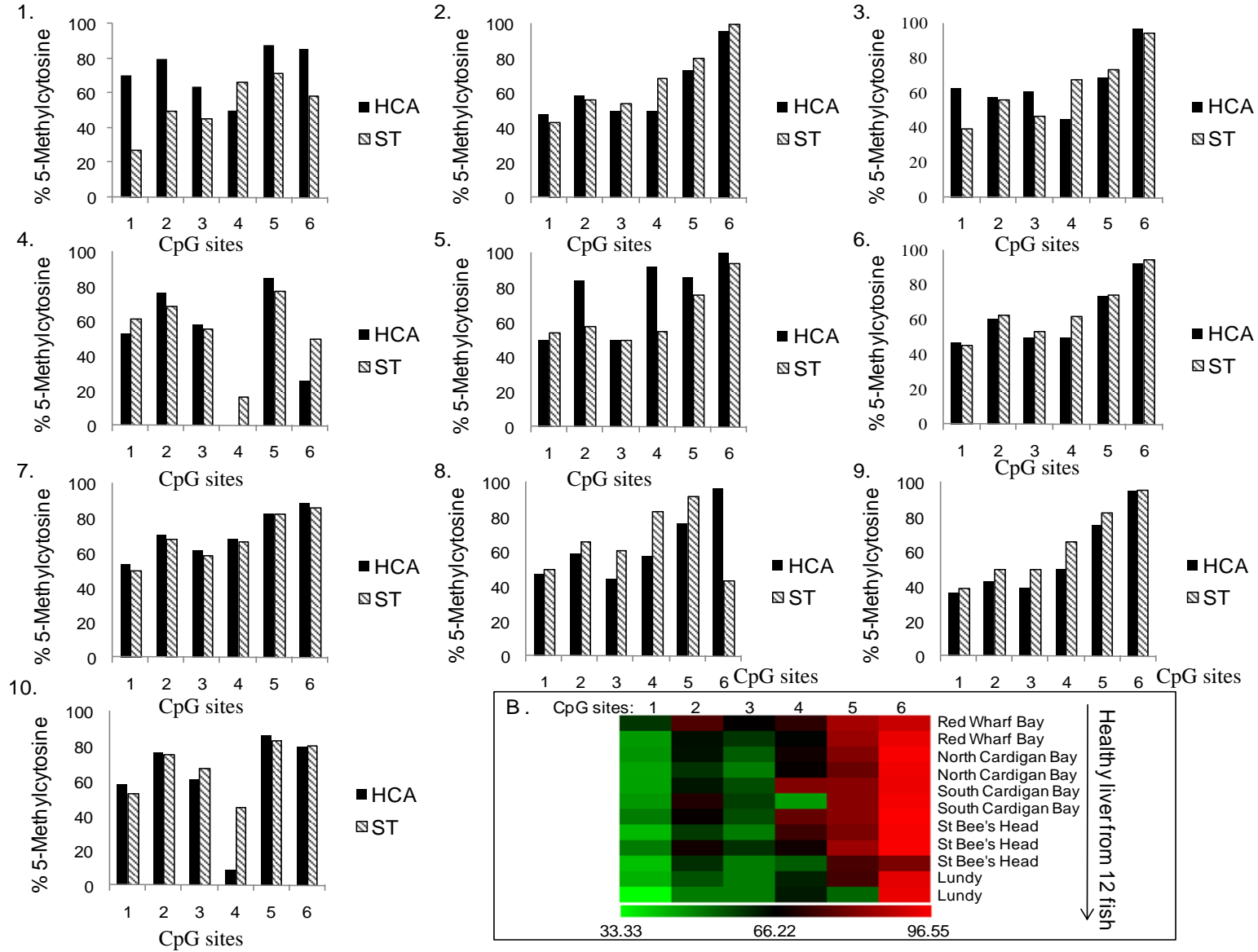
A. *protocadherin 1 gamma 22*



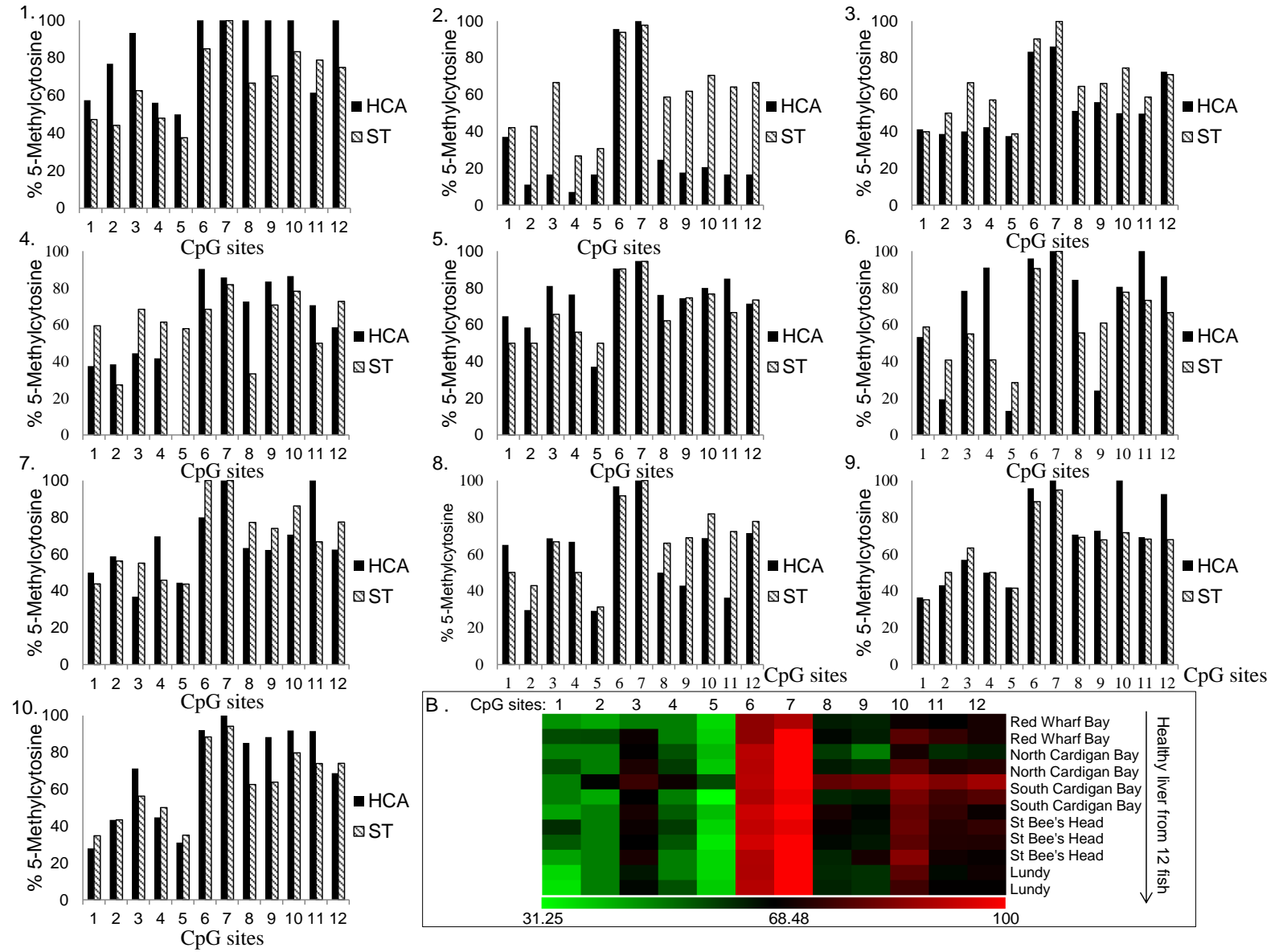
A. *microtubule-associated protein 1aa*



A. an unidentified sequence



A. *tubulin tyrosine ligase-like member 7*



A. Nidogen 1

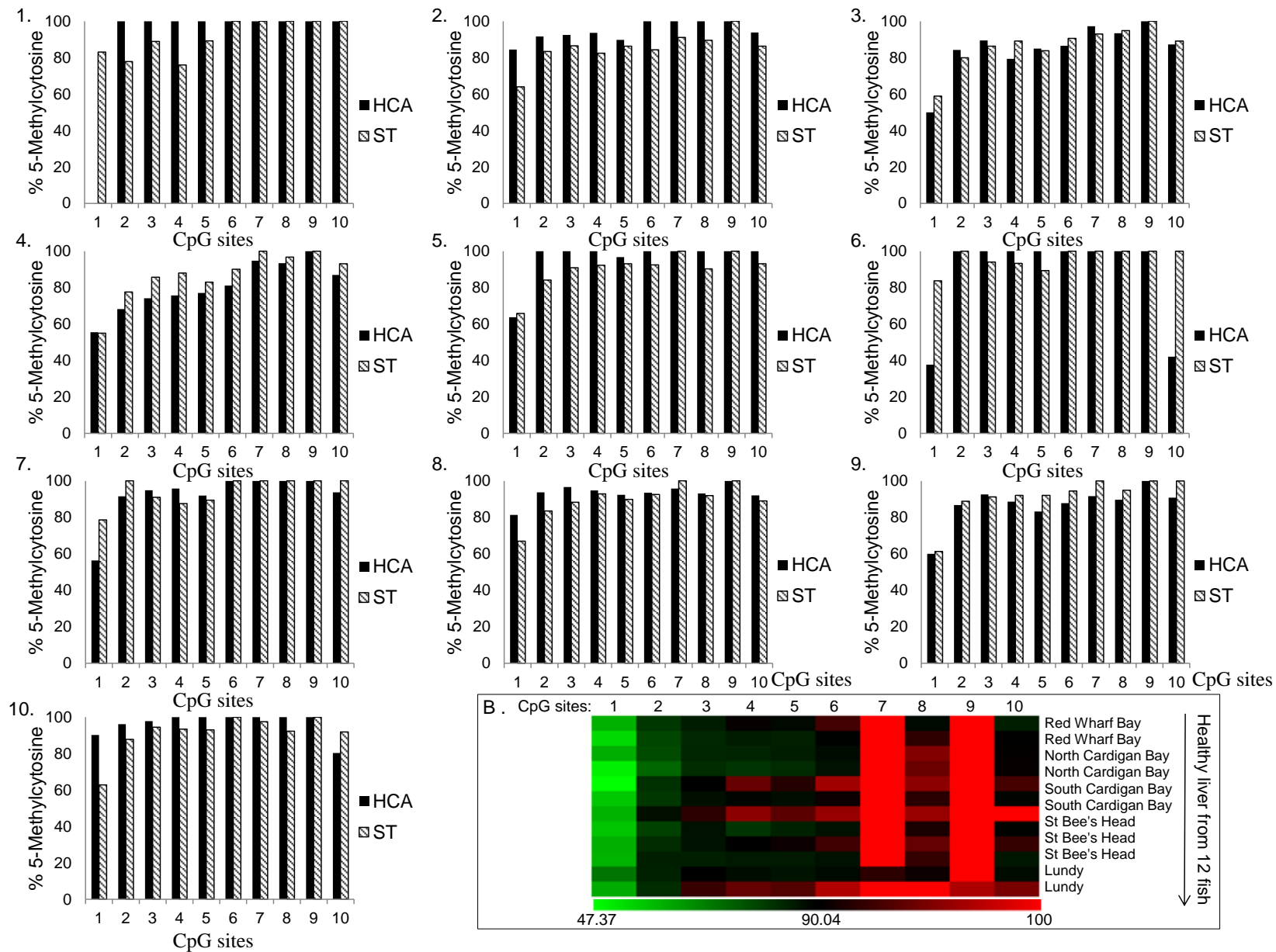


Figure 4.7. BSP was used to compare the methylation levels of *protocadherin 1 gamma 22*, *microtubule-associated protein 1aa*, an unidentified sequence, *tubulin tyrosine ligase-like member 7* and *nidogen 1* in dab collected from five different sampling sites in the Irish Sea and the Bristol Channel. Each page shows the result for one gene as indicated by the name of the gene at the top of the page.

Section A of each page: Methylation levels of CpG sites (x-axis) in 10 individual fish (10 HCA and 10 corresponding ST). Sampling sites: 1, 6, 9: South Cardigan Bay; 2: Inner Cardigan Bay; 3, 8: Red Wharf Bay; 4: North Cardigan Bay; 5, 10: St Bee's Head. Each graph represents one fish.

Section B of each page: Heat map of methylation levels for the same CpG sites in 12 healthy dab livers. For every analysed matrix each row represents one fish and each column represent a CpG site. Fish are labelled based on their sampling locations. The bar below the heat map represents the scale of methylation levels with green closest to lowest level of methylation, red closest to highest level of methylation and black for the CpG sites with median level of methylation.

4.3.2.3. Transcription profiling of dab tumours

4.3.2.3.1. Design of 8x15k gene expression microarray and gene expression analysis of dab hepatocellular adenoma

Dab DNA contigs derived from the MeDIP-HTS were annotated against the Genbank nr protein database for identification of coding sequences. The annotated sequences were used to design an Agilent custom-made dab specific 8x15k 60-mer oligonucleotide gene expression microarray as previously described (Chapter 2, section 2.5.3.3). Both the details of the designed microarray (name of the microarray: Agilent-031032_8x15k_BhamDab) and the experimental design (name of the experimental design: transcription profiling of dab hepatocellular adenoma liver) are available from ArrayExpress under accessions A-MEXP-2084 and E-MTAB-734, respectively.

RNA extracted from the 3 groups of interest whose DNA samples were used in the direct BSP experiment (HCA, ST and healthy dab liver) were labelled with fluorescent Cy5-CTP and hybridised to the dab custom-made gene expression microarray against a Cy3-CTP labelled RNA reference pool made from all samples. Following quality checks and normalisation, genes differentially expressed between the 3 groups investigated were identified (>1.5-fold change, FDR <15%, MIAME-compliant raw microarray data were submitted to ArrayExpress at EMBL-EBI and can be found under accession E-MTAB-734). During the final incubation of the slides in the oven, the hybridisation mixture on one of the slides containing an HCA sample evaporated. This was a manufacture's fault as the rubber seal between the chamber slide and the microarray slide was loose. Therefore, the data achieved from the HCA sample in question did not pass the initial quality checks and with its corresponding ST was removed from the analysis. Most differentially expressed genes were observed between HCA and ST with 520 genes statistically significantly altered in expression between HCA and ST samples (208 genes repressed and 312 genes induced in HCA compared to ST, paired *t*-test, >1.5-fold change, FDR <15%). This finding was not un-expected. Each pair of HCA and ST samples was dissected from the same fish. This reduced the levels of inter-individual variation that was observed when comparing liver tissues from healthy fish with HCA or ST samples from different fish. Therefore the main focus has been on the observed differences between HCA and ST. We found 91 genes to be statistically significantly altered in expression between HCA and healthy tissue with 73 and 18 induced and repressed in HCA compared to healthy

dab liver, respectively (>1.5 -fold change, FDR $<15\%$, unpaired t -test, Appendix File 12). Furthermore, there were apparent differences between ST and healthy dab liver. However, due to inter-individual variation they did not pass the FDR cut-off.

4.3.2.3.2. Principal components analysis

All the genes from the three categories of interest (i.e. disease) that were differentially expressed were subjected to principal component analysis (Figures 4.8, 4.9 and 4.10).

In all the PCA plots, samples were separated based on disease status with the clearest separations identified when comparing two sample groups at a time (Figures 4.8 and 4.9). As shown in Figure 4.8, PC1 and PC2 resulted in clear separation of HCA samples from ST samples with the first and second components accounting for 54.66% and 10.04% of variance, respectively. While PC1 and PC2 accounted for 58.14% and 7.3% of the variance, respectively in PCA plot of HCA and healthy dab liver samples.

In addition, within each condition samples were further clustered based on sampling sites. As shown in Figure 4.8, HCA and ST samples collected from Cardigan Bay were separated from other sampling sites along the PC1 axis. This demonstrates the influence of environmental factors upon the main condition (i.e. disease).

4.3.2.3.3. Hierarchical clustering

Hierarchical clustering (HAC) resulted in identification of distinct transcription profiles for HCA and ST and validated the PCA results (Figure 4.11). Hierarchical clustering, similar to PCA, is a statistical approach for identification of similar features within a data set. However, in contrast to the data reduction approach used in PCA, HAC relies on grouping of the similar features and generating a tree with all the data points present (Raychudhuri et al., 2001). As shown in Figure 4.11, a gene tree was generated in which genes with similar expression profiles and samples with similar transcription profiles were clustered together.

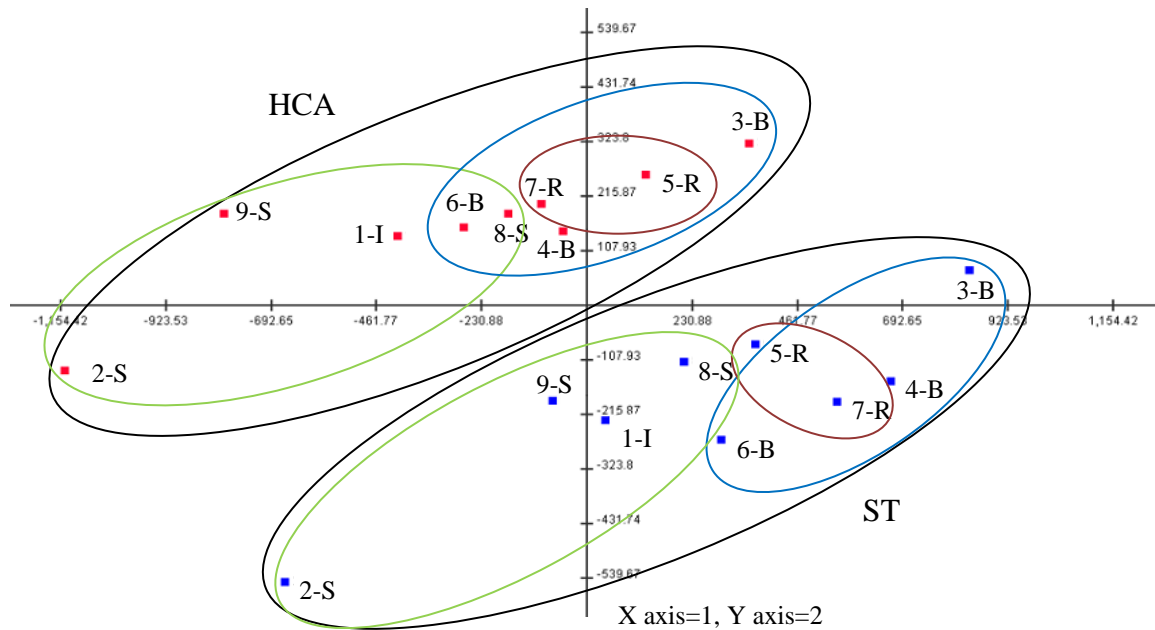


Figure 4.8. Principal component analysis (PCA) scores plot for all the genes that responded to the disease status of dab tissue samples. These genes were identified by the gene expression microarrays generated for HCA and ST. The plot shows separation of dab HCA samples (red) from dab ST (blue) samples based on treatment along the PC1 and PC2 axes. There is some level of separation based on sampling sites. Cardigan Bay is separated from the remaining sampling sites mainly along the PC1 axis. Numbers 1_9 represent 9 individual fish. Matching numbers are ST and HCA samples from the same fish. Sampling sites for each fish are also shown. I: Inner Cardigan Bay, S: South Cardigan Bay, B: St Bee's Head, R; Red Wharf Bay.

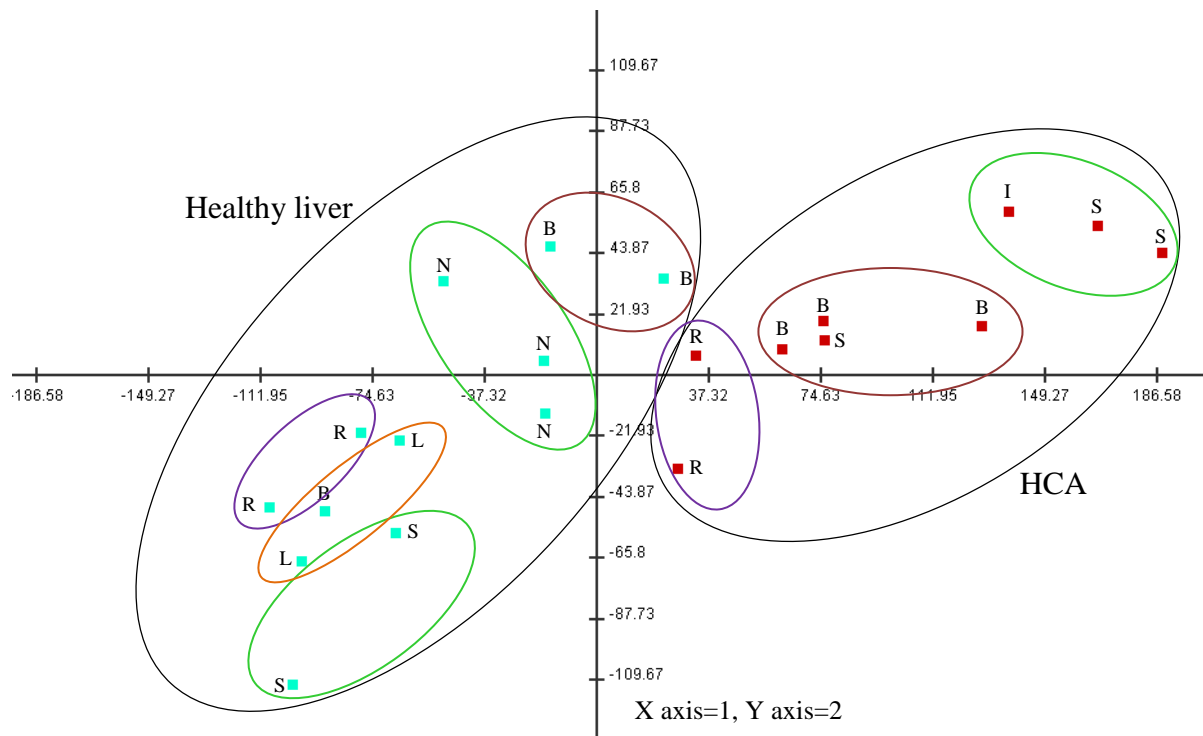


Figure 4.9. PCA scores plot for all the genes that responded to the disease status of dab tissue samples. These genes were identified by the gene expression microarrays generated for HCA and healthy tissue. The plot shows separation of the dab HCA samples (red) from healthy dab liver samples (green) based on treatment and along the PC1 and PC2 axes. In addition, there was clear clustering of samples based on sampling sites within each treatment group along the PC1 and PC2 axes. Dab healthy liver and HCA samples are represented based on their sampling sites. I: Inner Cardigan Bay, S: South Cardigan Bay, N: North Cardigan Bay, B: St Bee's Head, R; Red Wharf Bay, L: Lundy.

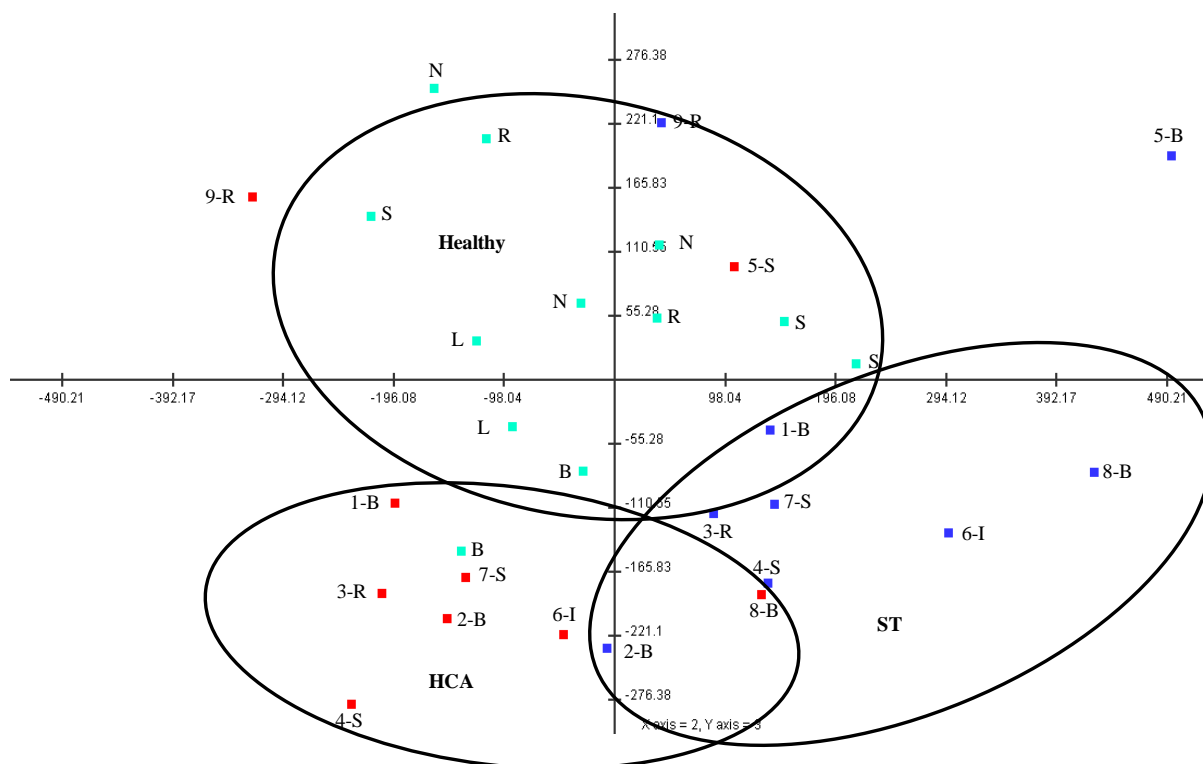


Figure 4.10. PCA scores plot for all the genes that responded to the disease status of dab tissue samples. These genes were identified by the gene expression microarrays generated for HCA, ST and healthy tissue. The plot shows a relatively good separation of dab HCA samples (red) from healthy dab liver samples (green) and ST samples (blue) based on treatment along the PC2 (10.04%) and PC3 (6.57%) axes. However no clear clustering of samples was observed based on sampling sites. Numbers 1_9 for HCA and ST represent 9 individual fish. Matching numbers are ST and HCA samples from the same fish. Sampling sites for each fish are also shown. I: Inner Cardigan Bay, S: South Cardigan Bay, N: North Cardigan Bay, B: St Bee's Head, R; Red Wharf Bay, L: Lundy.

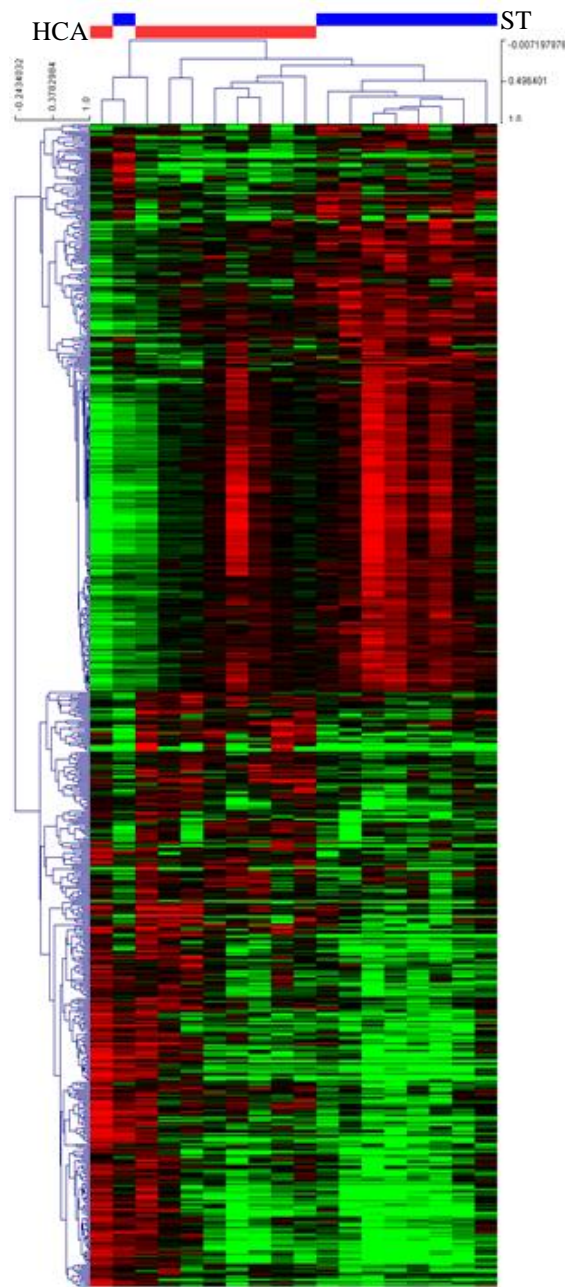


Figure 4.11. Hierarchical clustering tree generated using Pearson correlation showing relationships between differentially transcribed genes and disease status (HCA and ST). The tree is coloured based on a gradient with respect to transcription levels for each feature, with red indicating over-transcribed and green under-transcribed. Each column represents a specific sample and each row represents a specific gene. Blue bar: ST, Red bar: HCA. For clarity, individual genes have not been labelled.

4.3.2.3.4. Ingenuity Pathway Analysis

IPA was performed to characterise the functional relationships between genes with altered expression in HCA compared to healthy dab liver samples, as well as HCA and ST samples. IPA identified that the most significant networks and most enriched functions were commonly associated with cancer in both comparisons of HCA to ST and HCA to healthy dab liver. Networks and functions including cell death, cell cycle, cellular growth and proliferation, cell morphology, cell signalling, amino acid synthesis and gene expression and canonical pathways including molecular mechanisms of cancer, Wnt/ β -catenin signalling, aryl hydrocarbon receptor signalling and regulation of eIF4 and eIF2 signalling were associated with the differentially expressed genes for each comparison analysed (FDR <5%) as shown in Figures 4.12 and 4.13. In both sets of analysis (HCA and ST; HCA and healthy dab liver) cancer was highlighted as the top predicted disorder based on categories of the genes with altered transcription (FDR <5%, >1.5-fold change). As an example Table 4.4 shows a subsection of the genes (under- and over-expressed) associated with cancer derived from comparison of HCA to ST. The list of all the genes with altered transcription levels and their associated biological functions (FDR <5%) in both groups is shown in Appendix File 13. Figure 4.14 illustrates genes with altered expression levels in dab HCA compared to ST (>1.5-fold change, FDR <5%) related to the canonical pathway “molecular mechanisms of cancer” in human, including *v-akt murine thymoma viral oncogene homolog 1* (*akt1*), *cyclin d2* (*ccnd2*) and *mitogen-activated protein kinase 10* (*mapk10*).

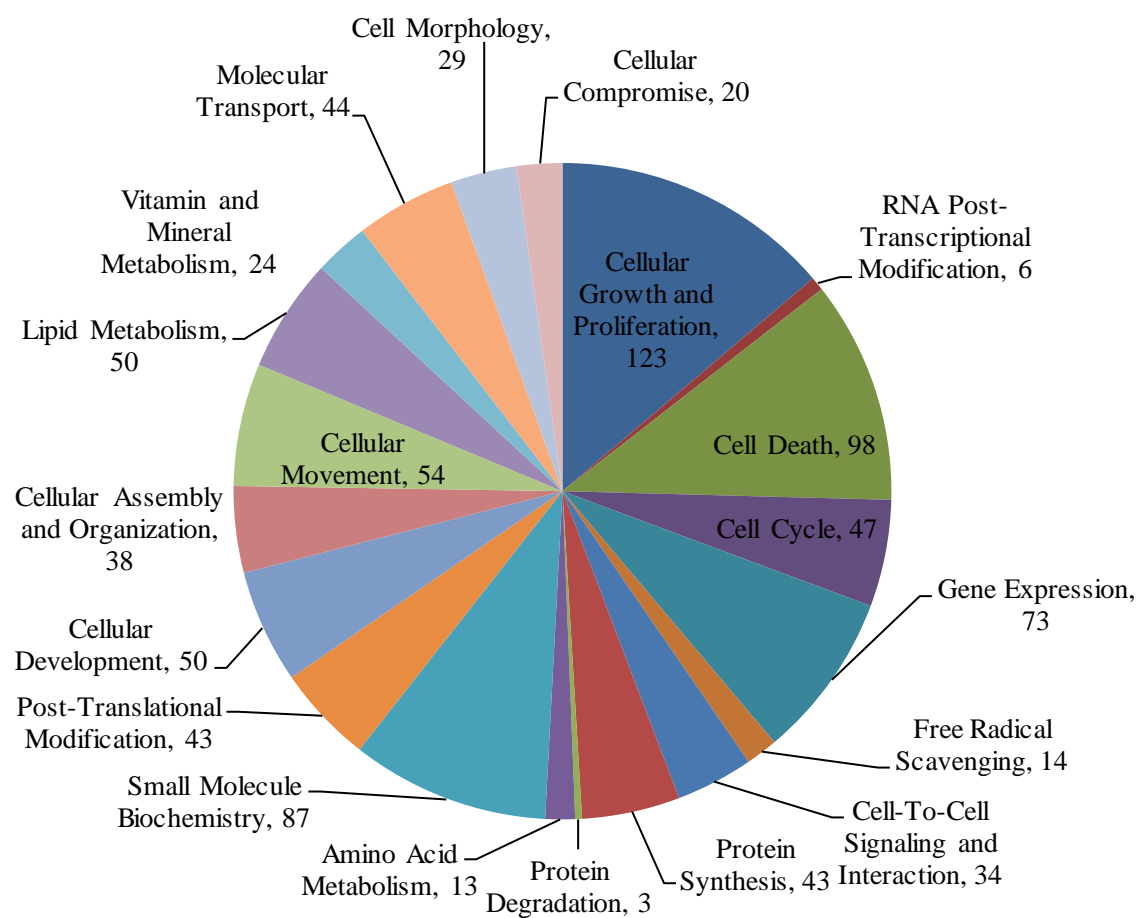


Figure 4.12. Biological functions enriched among genes with altered transcription levels (over- and under-expressed) in HCA compared to ST. IPA was used to group the genes with altered transcription levels based on biological functions (>1.5 -fold change, FDR $<5\%$). The numbers of genes associated with each biological function are shown.

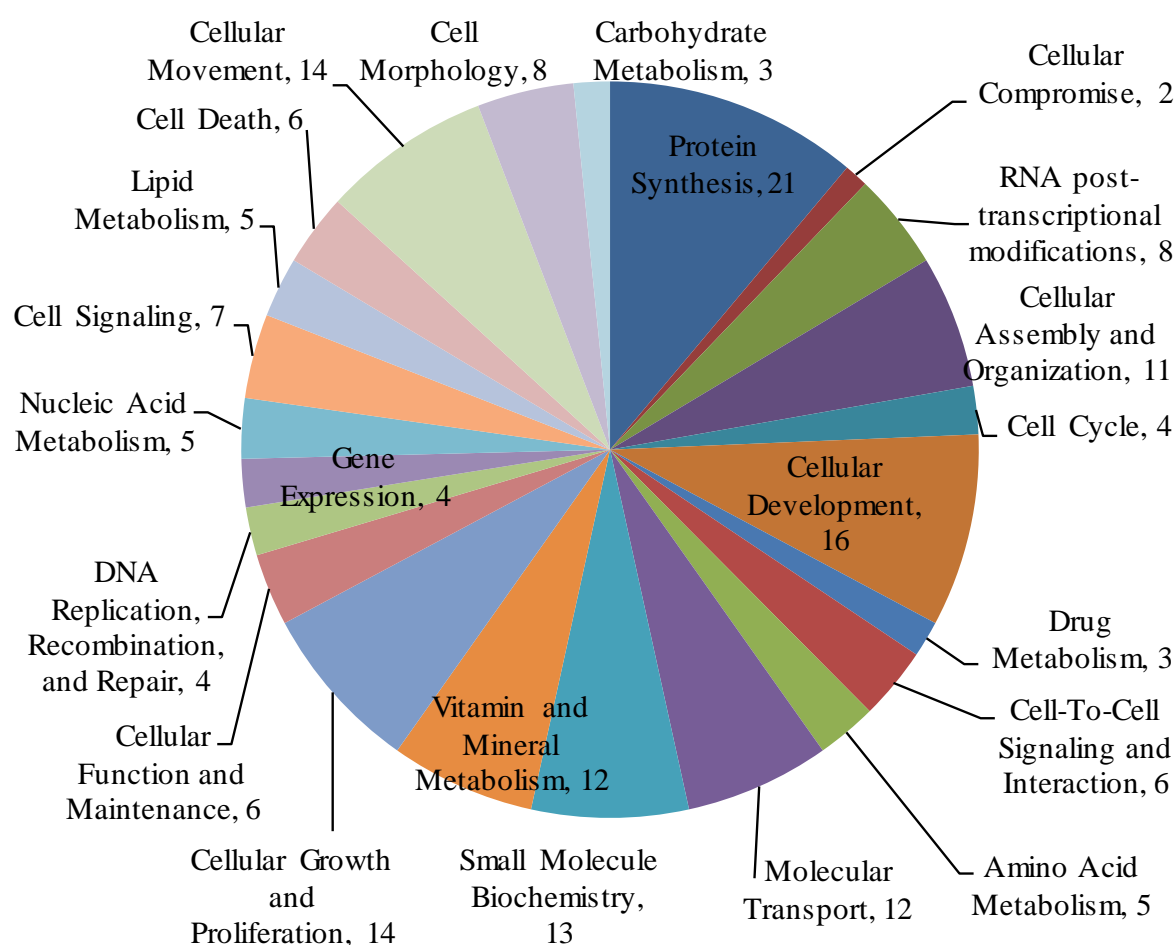


Figure 4.13. Biological functions enriched among genes with altered transcription levels (over- and under-expressed) in HCA compared to healthy dab liver. IPA was used to group the genes with altered transcription levels based on biological functions (>1.5 -fold change, $FDR < 5\%$). The numbers of genes associated with each biological function are shown.

Biological function	Gene expression level in HCA compared to ST
Cell death	<i>mitogen-activated protein kinase kinase kinase 10 (map3k10,↑), ras-like protein 1 (ras,↑), histone de-acetylase 1 (hdac1,↓), retinoid x receptor beta (rxrb,↑), cyclin D2 (ccnd2,↑), v-myc myelocytomatosis viral oncogene-like protein (myc,↑), dnaJ (hsp40) homolog, subfamily B, member 1 (dnajb1,↑)</i>
Cellular growth and proliferation	<i>S-adenosylhomocysteine hydrolase (ahcy,↑), retinoid x receptor beta (rxrb,↑), histone de-acetylase 1 (hdac1,↓), ras-like protein 1 (ras,↑), myc-associated zinc finger protein (maz,↑), cyclin d2 (ccnd2,↑), 26s protease regulatory subunit 8 (psmc5,↑), DnaJ (hsp40) homolog, subfamily B, member 1 (dnajb1,↑), suppression of tumorigenicity 14 (st14,↓)</i>
Cell cycle	<i>cell cycle regulator mat89bb homolog (m89bb,↓), mediator of RNA polymerase ii transcription subunit 29 (med29,↑), retinoid x receptor beta (rxrb,↑), histone de-acetylase 1 (hdac1,↓), cyclin d2 (ccnd2,↑)</i>
Protein and lipid synthesis and reproduction	<i>vitellogenin b (vtgb,↑), vitellogenin a (vtga,↑), 60s ribosomal protein l9 (rpl9,↑), 60s ribosomal protein l19 (rpl19,↑), 60s ribosomal protein l28 (rpl28,↑), aldehyde dehydrogenase 8 member a1 (aldh8a1,↑), Similar to human cytochrome P450, family 2, subfamily J, polypeptide 2 (cyp2j2,↑), eukaryotic translation initiation factor subunit 3 isoform cra_d (eif3,↑), methionine aminopeptidase 2 (metap2,↑)</i>

Table 4.4. Subsection of the genes with altered gene expression in HCA compared to ST. These genes are associated with the most relevant biological functions altered in tumours (FDR <5%, >1.5-fold change). ↑ and ↓ over- and under-expressed in HCA compared to ST, respectively. A complete list of the biological functions and associated genes is presented in Appendix File 13.

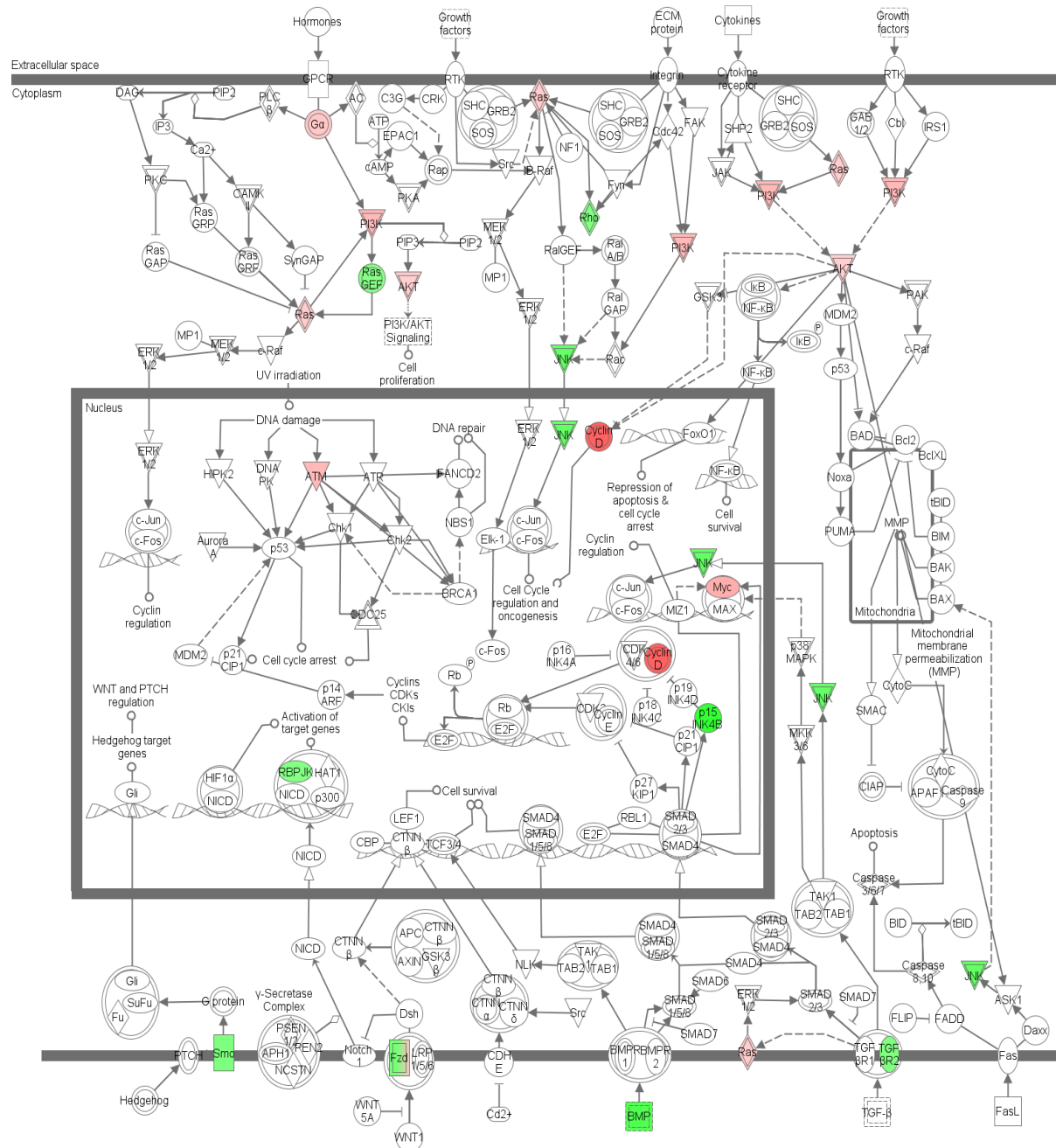


Figure 4.14. Biological network of genes linked to the canonical pathway “molecular mechanisms of cancer” with altered gene expression levels in dab HCA compared to ST (>1.5-fold change, FDR <5%). Green: under-expressed, red: over-expressed in HCA compared to ST.

4.3.2.3.5. Confirmation of the gene expression data using RT-PCR

Microarrays are extremely useful for achieving a comprehensive transcription profiling and identifying subset of genes with altered transcription between different experimental conditions. However, the experimental procedure is complex and the data can be influenced by each step. Hence, a second approach is commonly used to validate the data achieved from microarray experiments. Real time-PCR is recognised as an accurate and reproducible method for confirmation of the gene expression microarray data (Rajeevan et al., 2001). Therefore based on microarray data, four genes with statistically significantly altered transcription levels and acceptable primer efficiencies were selected for RT-PCR analysis. These genes included: *udp-glucose 6-dehydrogenase (ugdh)*, *vitellogenin c (vtgc)*, *vitellogenin b (vtgb)*, *S-adenosylhomocysteine hydrolase (sahh)* and *syntrophin, beta 2 (sntb2)*. As shown in Table 4.5 the gene expression data achieved from microarray and RT-PCR were in agreement. A statistically significant positive correlation was observed between microarray data and RT-PCR ($P < 0.05$, Pearson's correlation coefficient=0.97).

Gene	Grouping	RT-PCR fold change	Microarray fold change (FDR <15%, >1.5-fold change)
<i>vitellogenin b</i>	HCA and H	3.17**	2.13
<i>vitellogenin c</i>	HCA and ST	7.89*	8.5
<i>S-adenosylhomocysteine hydrolase</i>	HCA and ST	2.3*	1.78
<i>udp-glucose 6-dehydrogenase</i>	HCA and ST	2.51**	3.24

Table 4.5. Comparison of fold change in gene expression for the selected genes by RT-PCR and microarray analysis. The same mRNA was used for both RT-PCR and microarray analysis. *syntrophin*, *beta 2 (sntb2)* gene was used as a reference gene. Five biological and four technical replicates were used. (H: healthy dab liver, ** $P < 0.01$, * $P < 0.05$, P -value determined by 2-tailed Student's t -test following establishment of normal distribution of data and homogeneity of variance by Shapiro-Wilk test and Levenes' test, respectively).

4.4. Discussion

4.4.1. Global DNA methylation

Despite the observed differences in methylation levels of specific genes in HCA compared to the surrounding liver tissues, one of the most important findings of this study was the detection of similarly lower levels of global DNA methylation in both ST and HCA samples. A significant 1.8-fold global DNA hypomethylation was detected in HCA compared to liver from non-cancer bearing dab. A similar level of significant DNA hypomethylation was also identified when comparing “apparently healthy” liver (ST) surrounding the HCA tumours with liver from non-cancer bearing dab (Figure 4.2). Surprisingly, there were no detectable differences in global DNA methylation levels between ST and HCA dissected from the same liver.

There is evidence that global DNA hypomethylation and gene specific hypermethylation can occur during early stages of tumourigenesis, this has been reported in apparently normal tissues prior to mutational changes in patients who later developed cancer (Feinberg et al., 2006; Pogribny, 2010). It has been suggested, as part of the “progenitor model of tumourigenesis”, that epigenetic changes arise as early as progenitor cells. DNA hypomethylation and other gene specific epigenetic changes can lead to genome instability, chromosomal rearrangement, loss of imprinting and activation of germ line specific genes and transposons, therefore making gatekeeper genes more susceptible to mutations and increasing the risk of tumourigenesis (Gronbaek et al., 2007; Feinberg et al., 2006; Sharma et al., 2010; Pogribny, 2010; De Smet and Loriot, 2010). Mutations in gatekeeper genes of the pre-neoplastic cells and further epigenetic and genetic changes can ultimately lead to fully formed tumours (Sharma et al., 2010; Feinberg et al., 2006). Hence DNA hypomethylation is a striking feature and a common early mechanism in carcinogenesis (Feinberg et al., 2006). This is in agreement with our findings of DNA hypomethylation in apparently healthy portions of tumour-bearing dab livers (the epigenetic progenitor model of tumourigenesis and examples of detection of epigenetic alterations at early stages of tumourigenesis are described in Chapter 1, section 1.3.5.3).

4.4.2. Pilot study indicated change in the methylation of DNA samples extracted from HCA samples compared to ST at gene level

Based on the epigenetic model of tumourigenesis, it was anticipated that the “epigenetically primed” liver tissues would undergo further epigenetic and genetic changes, leading to formation of HCA tumours with specific methylation changes at individual genes (Sharma et al., 2010; Feinberg et al., 2006). Indeed the results of the pilot study, using MeDIP combined with a flounder cDNA microarray, showed that DNA methylation profiles of the investigated HCA samples differed from those of ST for specific genes. However, due to several reasons this study was only used as a preliminary experiment to test whether DNA methylation alterations could be detected at the gene level in HCA samples, and thereby justifying the need for a comprehensive DNA methylation study.

Due to sufficient levels of sequence homology between the two flatfish species, dab and flounder, the use of the flounder microarray has been successful for dab (Diab et al, 2008; Small et al., 2010; Cohen et al., 2007). However, a limited number of genes (3,336) were represented on this microarray, the probes were variable in length and targeted towards the 3' end of the genes. This excludes the epigenetically important control regions, such as the gene promoter regions near the transcription start sites and exon one regions. Thus, although this platform is useful for transcriptomic studies, it has limited use for comprehensive DNA methylation studies. Hence to compensate for this, longer DNA fragments were generated prior to MeDIP with the aim of detecting overlaps between long MeDIP fragments, potentially including gene promoters and the coding sequence amplicon microarray probes.

Nevertheless, due to the design of this microarray methylated fragments towards the 3' end were more represented on the microarray compared to the fragments located towards the 5' end of a gene. Therefore, the methylation status of the CpG dinucleotides located around the TSS and exon one, which are predicted to include the promoter regions and regulatory regions (Suzuki and Bird, 2008), are most likely under-represented in the results achieved from this cDNA microarray. Furthermore, as the methylation levels between different sections of a gene vary, it is important that each gene is represented with more than one probe tiling the entire region of interest. The latter will provide detailed information regarding the methylation profile of each section of the gene. In addition, dab samples used in this study were captured from different sampling sites located in the Irish Sea, North Sea and Bristol Channel. The

variability between different fish may be high due to different sampling sites. This can negatively impact the study by deflecting attention from the changes caused due to formation of tumours. Therefore to overcome this issue in the main study, in addition to increasing the number of samples, only liver samples from fish collected from the Irish Sea and Bristol Channel were used.

Regardless of the limitations of this pilot study, the PCA plot and the identified genes presented in Table 4.1 indicated that at the gene level HCA and ST samples differed in their methylation levels. For example, genes with biological functions such as angiogenesis, inflammation, antioxidant activity, metabolism, protein synthesis and drug transporters were identified to potentially have altered methylation levels in HCA samples compared to ST samples.

4.4.3. Overview of the DNA methylation changes

Based on the success of the pilot study, to achieve a comprehensive DNA methylation and transcriptomic profile for the un-sequenced flatfish dab, we successfully combined MeDIP with *de novo* HTS and designed a gene expression microarray based on the contigs achieved. As demonstrated for the first time by Down et al., (2008), MeDIP-HTS is a powerful, unbiased and quantitative technique for establishing whole genome methylation profiles in model organisms. However, this technique has never before been attempted for studying DNA methylation alterations in un-sequenced species.

Combining MeDIP with high-throughput sequencing resulted in considerable numbers of overlapping reads for construction of contigs, 34% of which were hypo- or hyper-methylated (>1.5-fold change) in tumour samples compared to control and could be annotated based on homology between protein-coding regions. This method is not without limitations. As the data are assembled based on conserved regions and homolog sequences, most of the identified regions corresponded to the CGIs located within exons, and only a fraction of the data focussed on promoter regions. Therefore the main outcome of this study was identification of genes and pathways with altered DNA methylation in HCA compared to ST irrespective of the position of the changes observed. Nevertheless, this approach resulted in identification of a substantial number of genes and pathways with altered methylation in HCA samples, as shown in Appendix Files 8, 9 and 10. Most importantly the data confirmed our hypothesis by

which we proposed that although similar in terms of global DNA methylation levels, HCA samples differ in DNA methylation at the gene level compared to primed surrounding tissue. Our finding that methylation of exons was different in the tumour samples, is in agreement with previous findings in humans (Liang et al., 1998). Furthermore, in tumourigenesis identifying DNA methylation alterations is more important than the absolute levels of DNA methylation (Liang et al., 1998). In humans, aberrant DNA methylation mediates changes in most of the key pathways associated with cancer, resulting in loss of cell cycle control, cell-to-cell signalling, apoptosis signalling, altered receptor and transcription factor functions as well as genomic instability (Baylin et al., 2001). Our study identified expression and DNA methylation changes in key genes of the same biological pathways in dab HCA samples. DNA methylation is therefore an important factor in tumourigenesis in the wild flatfish dab.

4.4.4. DNA methylation and transcriptional changes associated with tumourigenesis in dab HCA

Comparison of the gene expression data from the microarray experiment and DNA methylation data from the MeDIP-HTS experiment showed no significant overall inverse correlation between gene expression and DNA methylation levels (This data is presented in Appendix File 14). As discussed in Chapter 1 (section 1.3.2.3) and Chapter 3 (section 3.4.4) this is not un-expected. However, similar pathways were affected both in terms of DNA methylation and gene expression. For example, IPA identified changes in DNA methylation of genes involved in the Wnt/ β -catenin signalling pathway. Activation of this pathway has been associated with development of cancer mediated by stabilisation of β -catenin and inappropriate activation of β -catenin target genes such as *c-MYC* and *c-JUN* in humans (Polakis, 2000; Mann et al., 1999). Our data showed that methylation of genes such as *dishevelled*, *dsh homolog 2 (dvl2)* acting as a positive mediator of Wnt signalling, *frizzled* gene responsible for encoding Wnt receptors, tumour suppressor gene *adenomatous polyposis coli (apc)* and *c-myc* were altered in HCA compared to ST. In addition, transcription of *c-myc* and *frizzled homolog 7a* were induced and *frizzled homolog 10* and *7b* were suppressed in HCA samples.

Changes in methylation of genes involved in the apoptosis signalling pathway, involved in release of mitochondrial cytochrome c and activation of a range of caspases, were detected in

HCA samples. These genes included; *diablo* which is associated with neutralisation of members of inhibitory proteins (IPs) and release of cytochrome c (Adrian et al., 2001; Kohli et al., 2004), caspase activator *rock1* (Tsai and Wei, 2010; Coleman et al., 2001) and pro- and anti-apoptotic *nf- κ b* (Hu et al., 2000).

Based on transcriptional changes, the canonical pathway “molecular mechanism of cancer” and “cancer” as a disease annotation term were enriched among differentially expressed genes in HCA compared to ST and healthy liver. Genes involved in different aspects of tumourigenesis were significantly enriched in dab HCA samples compared to ST samples. For example, genes involved in cell cycle, proliferation and cell death (e.g. *cyclin d2* (*ccnd2*), *cyclin d binding protein 1* (*ccndbp1*), *ras* and MAP kinases (*mapk3k9*, *mapk10*)) were significantly over-expressed in tumours while genes related to repair mechanisms, such as *excision repair cross-complementing rodent repair deficiency, complementation group 6* (*ercc6*) and *suppression of tumourigenesis 14* (*st14*) were significantly down regulated. Genes associated with energy production and ribosomal protein synthesis were significantly over-expressed in tumours. This finding is not unexpected as both functions, like other parameters relating to proliferation/apoptosis as discussed above, are features likely to give tumour cells a selective advantage for clonal expansion and are required for highly proliferating tumour cells (Small et al., 2010).

4.4.5. Methionine cycle and DNA methylation

Changes detected in the expression levels of the genes involved in the methionine cycle between the three tissue categories of interest are suggestive of distortion of this pathway in dab HCA samples. During this cycle (Figure 4.15), methionine adenosyltransferase (MAT) catalyses the production of *S*-adenosylmethionine (SAM) from methionine. *S*-adenosylmethionine (SAM) acts as a universal methyl donor for a range of transmethylation reactions of biological significance, including regulation of gene expression by means of DNA methylation (Lieber and Packer, 2002). DNA methylation is catalysed by DNMTs. In this reaction a methyl group is transferred by DNMTs from SAM to the 5th carbon position of a cytosine base within the CpG dinucleotide, resulting in conversion of SAM to *S*-adenosylhomocysteine (SAH) (Gronbaek et al., 2007). In dab *DNA methyltransferase 3b* (-1.7 fold, matching the pufferfish *dnmt3b* gene from bases 6-116), *DNA methyltransferase 4* (-

2.55), *DNA methyltransferase 12* (-1.73) and *sam dependent methyltransferase* (-1.76) were hypomethylated in the tumour sample compared to ST. These changes suggest possible alterations in the expression of *dnmts*, the SAM to SAH metabolic pathway and subsequent change in DNA methylation.

Hyper- and hypo-methylated regions of DNA coexist in tumour cells. The possible mechanism behind this is explained in the demethylation-remethylation model and depends upon changes in the levels of DNMTs. As discussed in Chapter 1, section 1.3.5.1.1 and Figure 1.12, tumourigenesis involves an early transient phase of demethylation which will result in global DNA hypomethylation (Loriot et al., 2006). Several studies have shown that lower expression of *DNMT1* and *DNMT3b* causes severe DNA hypomethylation contributing to tumourigenesis as demonstrated in aggressive T-cell lymphomas in mice (Gaudet et al., 2003) and in human colorectal cancer cell line (Rhee et al., 2002). This is followed by a remethylation phase. In this stage, binding of transcription factors to un-methylated regions prevents them from becoming re-methylated. The newly established, lower DNA methylation level is maintained in tumour cells (Loriot et al., 2006; De Smet and Loriot, 2010). In fully formed tumours it has been shown that the expression levels of most DNMTs reach a similar or even higher level than in normal tissue as detected in various human cancers such as HCC and breast cancers (Girault et al., 2003; Sun et al., 1997; Laird and Jaenisch, 1994; Momparler and Bovenzi, 2000; Robertson et al., 1999). However, the contribution of different DNMTs to the maintenance of DNA methylation profile in different cancer cells vary and are highly dependent upon the relative abundance of different DNMTs in different tissue and cell types (Loriot et al., 2006). Therefore, it is possible that the expression profile of different DNMTs differs according to the type of tumour.

In addition, the *S-adenosylhomocysteine hydrolase* (*sahh*) gene was significantly over-expressed in dab tumour samples compared to ST. The conversion of SAH to homocysteine is catalysed by SAHH. SAH, due to its high binding affinity for the catalytic regions of SAM-dependent methyltransferases, is a potent inhibitor of transmethylation reactions (Lu, 2000; Lieber and Packer, 2002; James et al., 2002). SAH accumulation has been associated with global DNA hypomethylation and alterations in the expression of specific genes. Therefore constant hydrolysis of SAH via SAHH is necessary for maintaining transmethylation reactions (James et al., 2002). The observed difference in the expression level of the *sahh*

gene in dab tumour samples compared to ST samples indicates the possibility of alterations to the levels of SAH and homocysteine metabolites in dab HCA samples compared to ST samples.

The changes observed in the methylation and transcription of genes encoding enzymes of the methionine cycle led to the hypothesis that this pathway is affected during dab tumourigenesis. As a result, changes in the levels of the metabolites involved in the one-carbon cycle between the three categories of samples were investigated using LC-TSQ as shown in Chapter 5.

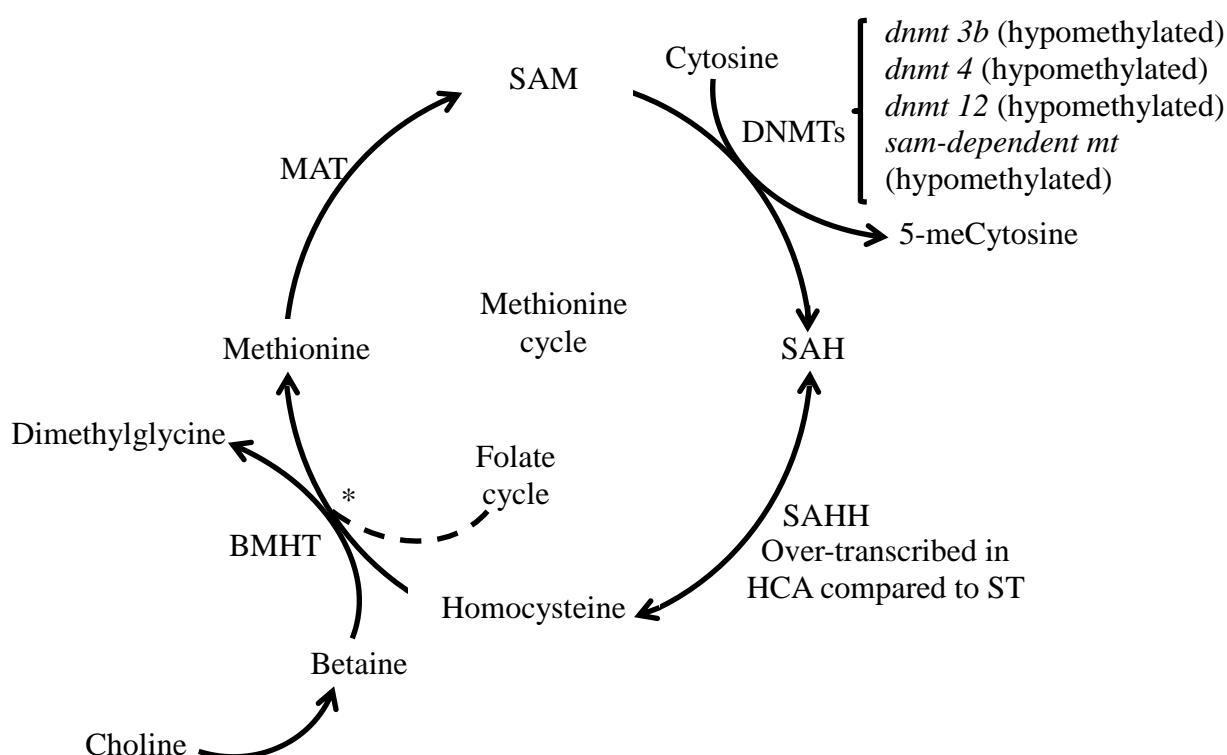


Figure 4.15. The methionine cycle. An illustration of the generation of the universal methyl donor, SAM from methionine. Transcriptional and DNA methylation alterations observed in this pathway in HCA samples compared to ST samples are indicated in this diagram. Abbreviations: SAM: *S*-adenosylmethionine; SAH: *S*-adenosylhomocysteine; Dnmt: DNA methyltransferase; SAHH: *S*-adenosylhomocysteine hydrolase; BMHT: betaine-homocysteine methyltransferase; MAT: methionine adenosyltransferase. Both folate and choline cycles are linked closely to this pathway via the site indicated by *. Folate donates one-carbon units for the production of methionine. Choline is located up-stream of betaine which is also used for the synthesis of methionine. Reduced levels of folate and choline inhibit the production of methionine and subsequently SAM (The one-carbon pathway is described in details in Chapter 1, section 1.3.3).

4.4.6. A link between environmental contaminants, changes in DNA methylation and transcription, and dab liver tumours

Statistically significant inductions in the expression of the egg yolk precursor proteins, *vitellogenins* (*vtga*, *vtgb* and *vtgc*), were observed in HCA compared to healthy liver. As this induction was also observed when comparing HCA to ST of the same fish, inter-individual variability, such as female reproductive cycle was not responsible. During oocytogenesis, vitellogenin proteins are transferred from the liver via the plasma to the growing oocyte where they are cleaved to form lipovitellin and phosvitin which serve as food reserves for growing embryos of most oviparous species (Zhang et al., 2011; Sumpter and Jobling, 1995). In contrast to high levels of vitellogenin in female fish during seasonal reproduction, very little vitellogenin is observed in male fish and juvenile female fish (Sumpter and Jobling, 1995). Expression of vitellogenin in the liver is controlled mainly through binding of estrogen receptors to estrogen response elements. Therefore, exposure to estrogen mimicking chemicals (endocrine disrupting chemicals (EDCs)) can elevate the expression of vitellogenin in fish, which can indicate feminisation of male fish. Hence, induction of vitellogenin in male fish and signs of intersex are used as biomarkers of exposure to EDC (Scott et al., 2007). Our findings agree with previous studies demonstrating induction of vtg in female and male dab and intersex in male dab (Small et al., 2010; Stentiford and Feist, 2005; Scott et al., 2007).

Different types of PCBs, polybrominated diphenyl ethers (PBDEs) and heavy metals such as cadmium (Cd) and lead (Pb) were detected in dab tissue. The concentrations of these chemicals, as shown in Appendix File 15, varied at different sampling sites with Cardigan Bay, specifically South Cardigan Bay, identified as the most polluted site in regards to PCBs and PBDEs investigated in dab liver tissues (Chemical measurements were conducted by the Centre for Environment, Fisheries and Aquaculture Science (Cefas, Weymouth, UK)). For example, from the seven non-planar PCBs recommended for monitoring by the International Council for the Exploration of the Sea (ICES), the concentration of PCB118 in dab liver at South Cardigan Bay was higher than the recommended hazardous ecotoxicological assessment criteria (EAC). This indicates the possibility of chronic exposure of dab to this chemical which is in agreement with the data reported by the Department for Environment Food and Rural Affairs (DEFRA) (UK marine monitoring and assessment strategy community (UKMMAS), Charting Progress 2 Feeder report, 2010). The differences between

sampling sites based on chemical data were also reflected in the PCA plots of gene expression data. As shown in Figures 4.8 and 4.9, samples in each disease category were further clustered based on sampling sites. Previous chemical exposure studies in European flounder, conducted in our laboratory, demonstrated that gene expression profiles alter distinctively following exposure to environmentally relevant concentrations of chemical contaminants, including cadmium and 17 β -estradiol (Williams et al., 2006; Williams et al., 2007). It is therefore possible that environmental contaminants are similarly influencing transcriptional profiles of flatfish dab. In addition, it has been reported that vitellogenin levels in male dab collected at different sampling sites varies, with fish collected from Cardigan Bay having the highest levels of vitellogenin compared to fish collected from other sampling sites investigated in this study (St Bee's Head, Red Wharf Bay and Lundy) (Scott et al., 2007). Furthermore, a study by Lyons et al., (2006) reported that dab collected from Cardigan Bay, specifically South Cardigan Bay, have higher tumour prevalence compared to dab collected from Red Wharf Bay. As the levels of environmental contaminants vary between sampling sites, it is plausible that environmentally driven epigenomic changes are responsible for different levels of susceptibilities of these fish. The main estrogenic compounds from the sewage treatment (i.e. the synthetic hormone 17 α -ethinylestradiol used in contraceptive pills and 17 β -estradiol) are short-lived and unlikely to reach so far off shore. Thus it is unlikely that these compounds would reach fish populations in offshore waters and cause adverse health effects such as elevated VTG levels (Scott et al., 2007). Hence, it is hypothesised that a widespread mixture of weak but persistent lipophilic estrogenic chemicals such as PCBs, PBDEs and some of their metabolites (Siddiqi et al., 2003) are responsible for the observed increase in expression of vitellogenin and potentially contribute to formation of these tumours (Scott et al., 2007).

A link between exposure to chemicals such as EDCs (PCB, EE2, E2) and heavy metals (e.g. As, Cd), changes in DNA methylation (Vandeghechuchte and Janssen, 2011) and tumourigenesis in rodents and humans has been established (Arita and Costa, 2009; Li et al., 2003; Lo and Sukumar, 2008; Skinner et al., 2010; Baccarelli and Bollati, 2009). For example, Cd can bind to DNA methyltransferases and inhibit their DNA methylation activity. This leads to hypomethylation and activation of genes, such as oncogenes and can lead to carcinogenesis (Baccarelli and Bollati, 2009). Although recent studies have demonstrated that EDCs can cause changes in DNA methylation in aquatic species (Stromqvist et al., 2010; Aniagu et al., 2008), with separate studies suggesting that EDCs can cause tumours in fish

liver (Tilton et al., 2006), this connection between exposure to EDCs, change in DNA methylation and tumourigenesis has not been established in fish.

Methylation and/or expression of several genes in both non-genomic and genomic (estrogen response element (ERE)-dependent and ERE-independent) estrogenic pathways were altered in HCA samples compared to healthy liver and ST samples. These genes are involved in biological functions such as proliferation and growth, differentiation, cell cycle, apoptosis and oncogenesis (Obrero et al., 2002; Safe and Kim, 2008; for a review see Bjornstrom and Sjoberg, 2005). The methylation and/or expression of several genes controlled through the ERE-dependent pathway that are commonly implicated in cancers were altered in this study. For example, induction in transcription of *c-myc* oncogene, *vtg a* and *vtg b*, reduced transcription of protease *cathepsin (cts)* with change in methylation of *c-myc* and *cts* were identified in HCA compared to ST (Obrero et al., 2002; Bjornstrom and Sjoberg, 2005). The estrogen-estrogen receptor (E-ER) complex can also modulate transcription of a range of genes without directly binding to DNA. This ERE-independent mechanism occurs in the nucleus and through protein-protein interaction with other DNA binding transcription factors (TF) and subsequent binding of TFs to their recognition sites in the promoter regions (Bjornstrom and Sjoberg, 2005). Furthermore, methylation/expression of several transcription factors that bind to ER and their target genes were altered in HCA compared to ST, such as *specificity protein 1 (Sp1)* involved in crucial biological processes (such as differentiation, proliferation, apoptosis) which is regularly over-expressed in tumours (Safe and Kim, 2008), *signal transducer and activator of transcription (stat) 5*, *activating transcription factor 2 (atf-2)* and *nuclear factor nf-kb* associated with inflammation and oncogenesis (Dolcet et al., 2005) and their target genes such as *c-myc*, *vtgs* and *cyclin d1* (Bjornstrom and Sjoberg, 2005). In addition to the changes observed in ERE-dependent and independent pathways, methylation and/or expression of several genes involved in the non-genomic estrogenic mechanism of action and development of cancer were also changed. For example, *insulin growth factor-1 receptor (igflr)*, known to be altered in terms of DNA methylation and expression in various cancers (Mirbahai et al., 2011), several different MAP kinases and oncogene *Akt kinase*, involved in cell proliferation and survival of tumour cells (Rosen and She, 2006), were altered in HCA samples. In the non-genomic pathway of estrogenic compounds, estrogen binds to the ER located at the cytoplasmic membrane and in combination with several other signalling molecules activates MAP kinases. Activated

kinases regulate transcription factors associated with ER (such as SP1, ATF2, STAT5 and NF-κB) as well as the nuclear ER through phosphorylation (Bjornstrom and Sjoberg, 2005). Figure 4.16, adapted from Bjornstrom and Sjoberg (2005), demonstrates the 3 main estrogen receptor pathways and the changes detected in dab HCA samples.

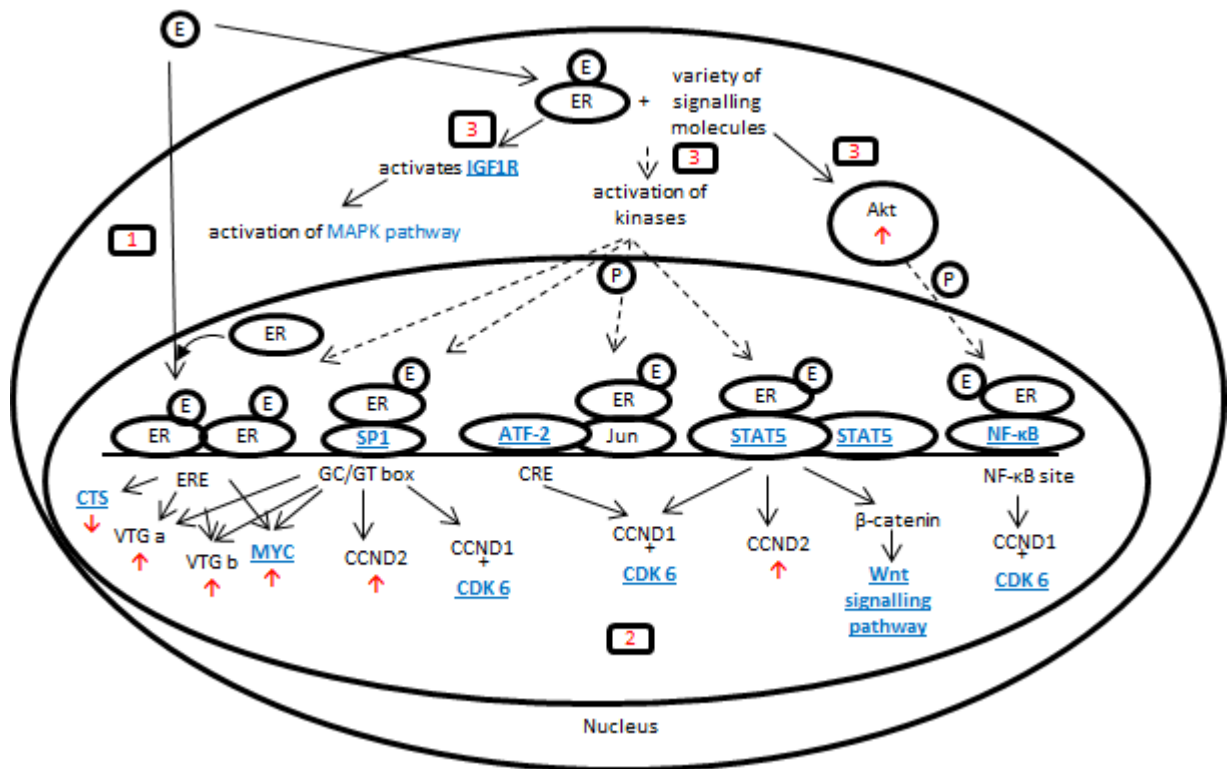


Figure 4.16. DNA methylation and transcription changes in genomic and non-genomic estrogenic pathway in HCA samples. 1. The classical genomic ERE-dependent pathway. 2. Genomic ERE-independent pathway. ER modulates the transcription of genes through other transcription factors such as SP1, ATF1, STAT5 and NF-κB. 3. Nongenomic pathway. Activation of the cytoplasmic ER subsequently leads to activation of kinases which regulate the function of TFs (e.g. SP1, STAT5 and NF-κB) and estrogen receptor through phosphorylation. Therefore all pathways are linked and influenced by each other. Many of the targeted genes are involved in key functions such as proliferation and cell cycle. Blue: genes with altered methylation in HCA compared to ST. ↑↓: genes with altered transcription (induced, repressed) in HCA compared to ST or healthy dab liver. P: Phosphorylation. E: estrogen or estrogen mimicking compounds. Adapted from original figure of Bjornstrom and Sjoberg (2005).

4.4.7. Hypothesis

Based on the evidence discussed above (chemical data and vitellogenin induction), one possible hypothesis is that chronic exposure to a mixture of contaminants such as endocrine disrupting chemicals (e.g. PCBs, PBDE) or metals (e.g. Cd) causes the initial changes in overall DNA methylation levels seen in the entire liver of tumour bearing dab. Indeed, inhibition of the activity of DNA methyltransferases is a reported effect of Cd (Baccarelli and Bollati, 2009). Exposure to EDCs and metals, especially during critical stages such as embryogenesis and gametogenesis, as well as increasing the individual's susceptibility to disease, can cause heritable epigenetic changes (Baccarelli and Bollati, 2009). It is plausible that these chemicals further affect the ERE-dependent and independent genomic and non-genomic estrogen response pathways. Changes in the methylation of genes in this pathway can alter the expression of genes involved in proliferation and cell cycle (Figure 4.16). It is postulated that additional epigenetic and genetic changes of the primed pre-neoplastic cells, and genome instability caused by global DNA hypomethylation, could lead to development of these tumours (Figure 4.17, adapted from Feinberg et al., 2006).

Although the combined data on gene expression, DNA methylation, chemical exposures and evidence available in the literature are in agreement with this interpretation, it is also possible that altered DNA methylation and disturbance of estrogen signalling pathway in ST is a secondary change due to proximity of this tissue to HCA in the organ. However, there is no evidence in relation to fish tumours to support this possible explanation. Further support for a direct environmental influence on DNA methylation changes and to test for tissue specificity might come from analysis of additional tissues in animals with modified DNA methylation in the liver.

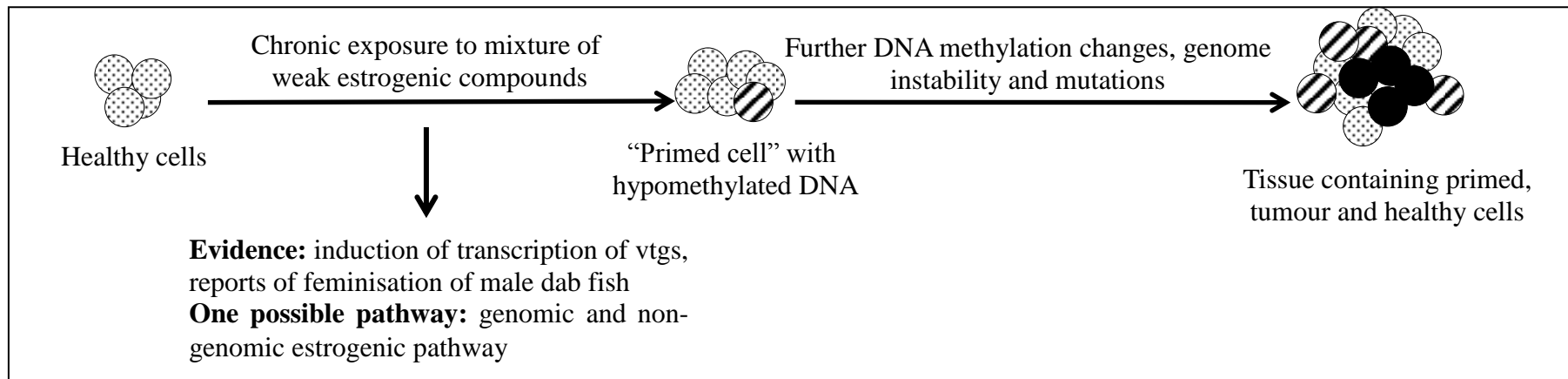


Figure 4.17. A possible mechanism of tumourigenesis in dab. Chronic exposure to a mixture of contaminants, such as weak estrogenic chemicals causes the initial global DNA hypomethylation. Over time, a fraction of the primed cells will further undergo epigenetic and genetic changes which may lead to formation of tumours. This hypothesis is based on the observed changes in DNA methylation, gene expression and chemical data. However, as their environment is complex most certainly there are other factors (chemical and biological) involved in the formation of these tumours. ●: Healthy cell; ▨: “Primed cell”; ●: tumour cell

4.4.8. Conclusions

In this study we demonstrated the suitability of MeDIP- *de novo* HTS for studying DNA methylation in un-sequenced species. We also showed for the first time, a link between DNA methylation abnormalities and liver tumours in wild fish. In addition, we proposed a mechanism to explain formation of dab liver tumours that arise in wild fish chronically-exposed to contaminants, such as EDCs. It was shown that DNA methylation and transcription of cancer-related pathways were altered, with inter-individual variation, in HCA compared to ST samples as demonstrated by BSP. Our data support the epigenetic model of tumourigenesis in which DNA methylation changes initiate tumourigenesis and therefore have the potential to be used for identification of tumour bearing livers, possibly even prior to manifestation of microscopic tumours. In particular, global DNA hypomethylation and the methylation status of DNA methyltransferases and the genes associated with the non-genomic and genomic estrogenic pathways have the potential to be used as early biomarkers of environmentally-induced epigenetic change, thus acting as indicators for identification of potential tumour inducing, hazardous environments for marine species. However, further studies are required to investigate if the observed transcription and DNA methylation changes are consequences or causes of these tumours. It is recognised that the epigenetic changes seen in ST are likely to be in association with potential genetic toxicity (e.g. PAHs). Nevertheless, it appears that there is a link between environment and dab tumourigenesis as the prevalence of these tumours differed between sampling sites. Overall, our study highlights the importance of incorporating and considering epigenetic mechanisms, especially in chronic exposures, for establishing acceptable levels of contaminants in marine environments which may subsequently impact the health of organisms, including humans.

Chapter 5

Changes to the one-carbon cycle in dab liver tumours

5.1. Introduction

Normal cells display a finely tuned epigenomic equilibrium (Esteller and Herman, 2002) regulated by a range of enzymes (Turner, 2009). However, as discussed in Chapters 1 and 4, the epigenome is sensitive to environmental factors (Jirtle and Skinner, 2007). In higher eukaryotes the immediate environment influencing cells is determined by the metabolism and physiology of both neighbouring cells and organs. The activities of enzymes modifying the epigenome, similar to most other enzymes, depend upon levels of cofactors and metabolites present in the intra- and extra-cellular environment (Turner, 2009). Therefore, the epigenome can be altered not only by environmental factors, such as exposure to chemicals like cadmium (Baccarelli and Bollati, 2009), but also by changes in the levels of cofactors and metabolites required for the activity of these enzymes (Turner, 2009; Herceg and Vaissiere, 2011). The interaction between methylome, transcriptome and metabolome and the influence of environmental factors at multiple levels is illustrated in Chapter 1, Figure 1.14 and Chapter 6, Figure 6.1.

The one-carbon cycle is the main biochemical pathway regulating DNA methylation by altering the production of the immediate methyl donor, *S*-adenosylmethionine (SAM), through feedback regulatory mechanisms. In this pathway folate, choline and methionine are the primary dietary methyl donors (Herceg and Vaissiere, 2011). Following several chemical reactions (described in detail in Chapter 1, section 1.3.3) SAM is produced as the final methyl donor. SAM is used in several transmethylation reactions, such as DNA methylation (Krijt et al., 2009) and it is second only to ATP in the variety of reactions in which it participates (Kumer et al., 2008). DNA methylation is catalysed by DNA methyltransferases (DNMTs). DNMTs transfer a methyl group from SAM to DNA, leading to conversion of SAM to *S*-adenosylhomocysteine (SAH) (Kumar et al., 2008). SAH is a potent inhibitor of transmethylation reactions. Accumulation of SAH, independent of SAM levels, inhibits DNA methylation reactions. Therefore, continuous removal of SAH is critical in regulation and maintenance of DNA methylation reactions (Caudill et al., 2001; James et al., 2002). Indeed, an increase in SAH levels and subsequent change in the SAM/SAH ratio has been associated with a range of clinical conditions (Krijt et al., 2009). Therefore, changes in this pathway (i.e. change in the transcription levels, the activity of the enzymes involved in this pathway or the levels of the metabolites) can influence DNA methylation. For example several studies in

rodents have demonstrated that diets deficient in primary methyl donors such as choline, folate and methionine can lead to development of tumours, especially in the liver. These tumours are induced through alterations of several biological pathways including the one-carbon cycle and consequently DNA methylation levels (Poirier et al., 2001; Ghoshal and Farber, 1984; de Camargo et al., 1985).

Previously (Chapter 4, section 4.4.5), it was shown that the transcription levels and DNA methylation of the genes involved in the methionine cycle, part of the one-carbon cycle, were altered in dab HCA tumours. The significance of this pathway in regulation of DNA methylation and tumourigenesis in mammals is evident (Krijt et al., 2009; Poirier et al., 2001; Ghoshal and Farber, 1984). Therefore, the main aim of the experiments described in this chapter was to conduct a targeted metabolomics study to investigate the differences in levels of the key metabolites involved in the one-carbon cycle between the three categories of interest, HCA, ST and healthy dab liver tissues.

Metabolomics is the investigation of the low molecular weight metabolites within a cell, tissue or biofluid (Viant, 2007). Several metabolomics methods have been used to measure the levels of SAM and SAH and other metabolites involved in the one-carbon cycle. These include liquid chromatography (LC) methods with electrochemical detection (Yi et al., 2000; Melnyk et al., 2000), LC methods with fluorescence detection after conversion of the analytes to fluorescent analogues (Davis et al., 2005; Kohlmeier et al., 2005) and more recently LC coupled to an ionisation source attached to a triple quadrupole tandem mass spectrometric (MS/MS) detection method (Struys et al., 2000; Krijt et al., 2009; Jiang et al., 2009; Liang et al., 2009). The latter method is an extremely powerful and sensitive technique for detection and quantification of specified compounds (Liang et al., 2009; Cox et al., 2005; Lujan et al., 2008). During LC a combination of mobile phases (e.g. water, acetonitrile, methanol), chemical modifiers (e.g. ammonium formate, ammonium acetate) and chromatography columns are used to achieve separation of the analytes of a complex matrix. Analytes are ionised at the ion source under a combination of high voltage and heat (Korfmacher, 2005) and directed to the mass analyser. In the triple quadrupole mass analyser, data acquisition is performed as multiple reaction monitoring (MRM) which creates the high sensitivity and selectivity of this approach (Liang et al., 2009). MRM is a method comprised of two ion selection stages (quadrupoles (q) 1 and 3) and a fragmentation ion stage (q2) in between the

two selection stages. The fragmentation of the parent ions is achieved by collision with gas atoms. Therefore, as a specific parent ion and specific fragment ions of the parent ion are selected for detection, this method is very precise. It is highly unlikely that two analytes would have the exact pattern of fragmentation and parent mass to charge ratio (m/z), making this approach ideal for analysis of complex biological matrixes (Figure 5.1) (Cox et al., 2005; Anderson and Hunter, 2006). Therefore, in this chapter LC combined with a triple stage quadrupole (TSQ) tandem mass spectrometer (MS/MS) were used to investigate the levels of key one-carbon metabolites in the three categories of interest, HCA, ST and healthy dab liver samples. Furthermore, DNA methylation and gene expression data related to the one-carbon cycle from the previous chapter were combined with the data from the metabolomics study conducted in this chapter to achieve a better understanding of the role of the one-carbon pathway in the development of dab HCA tumours.

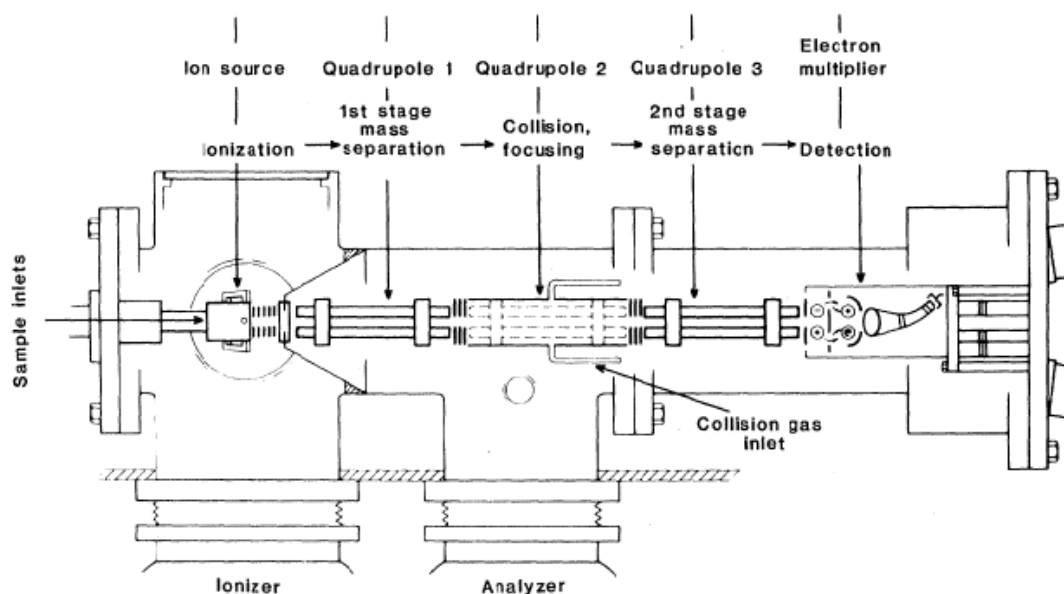


Figure 5.1. A schematic representation of an ion source coupled to triple quadrupole tandem mass spectrometer. In the ion source the analytes are ionised under a high heat and voltage. In quadrupole 1 (q1) or first stage selection, a specific ion of a known m/z is selected. The selected ion is fragmented in q2. In q3 the product ions with specific m/z that were generated from the fragmentation of the specified parent ion are selected (second stage selection). Finally, the selected parent and product ions are detected and recorded (This figure was modified from McLafferty, 1981).

5.2. Overview of the experimental approach (Methodological details are described in Chapter 2)

The dab liver samples used in Chapter 4 were also used for the quantitative targeted metabolomics study of the one-carbon cycle. The details of the samples are presented in Chapter 2, Table 2.1. DNA methylation, gene expression and relative quantification of selected metabolites were therefore investigated in the same 32 dab liver samples (10 HCA, 10 ST and 12 dab healthy liver). Metabolites were extracted using a modified version of multi-sample ceramic bead-based system described by Southam et al., (2008). In this method, an aliquot of the homogenised tissue (no less than 50 μ l, equivalent to 5 mg of wet tissue mass, ideal amount 10 mg (100 μ l)) as described in Chapter 2, section 2.2.2.3.2 was used for the extraction of metabolites. As the metabolites investigated were water soluble, only the polar fractions were used in this study. An aliquot of the prepared samples equivalent to 4.5 mg of tissue and the appropriate chemical standards were analysed with liquid chromatography (LC)-triple stage quadrupole (TSQ) mass spectrometry as described in Chapter 2, section 2.8.

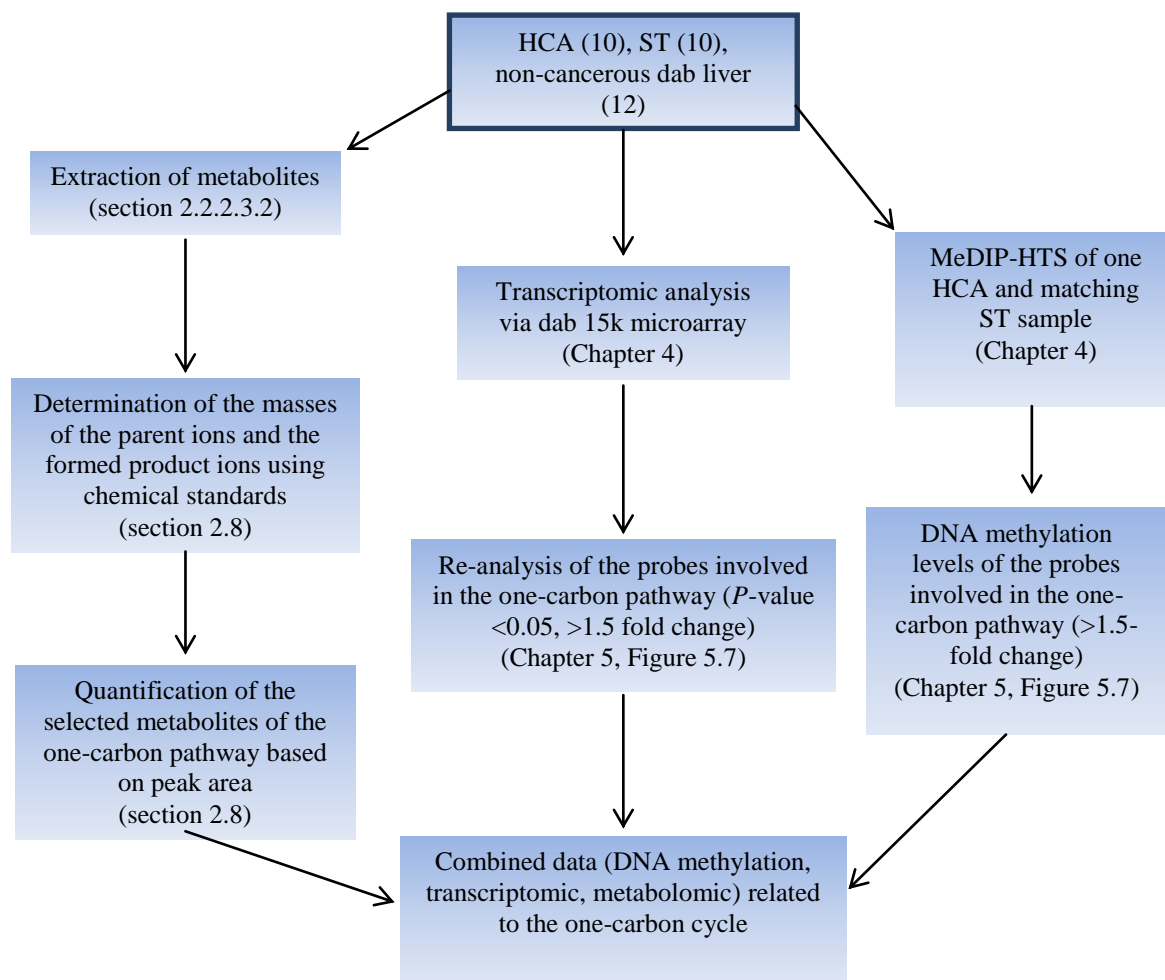


Figure 5. 2. Flowchart of the experimental procedures used in Chapter 5. The starting point of the diagram is the box with borders. The methods used in this chapter are described in detail in Chapter 2. In this flowchart, the sections where these methods are described in Chapter 2 are shown in brackets.

5.3. Results

5.3.1. Targeted quantification of 12 metabolites involved in the one-carbon cycle using LC-TSQ

The differences in the levels of 12 specific metabolites involved in the one-carbon cycle in the three categories of interest, HCA, ST and healthy dab liver samples, were investigated using LC-TSQ as described in Chapter 2, section 2.8. The metabolites quantified included: *S*-adenosylhomocysteine (SAH), *S*-adenosyl-*L*-methionine (SAM), methionine, folic acid, adenosine, betaine, sarcosine, homocysteine, glycine, dimethylglycine, choline and stachydrine (proline-betaine or dimethylproline).

As described in Chapter 2, section 2.8, analytes were detected with data acquisition performed as MRM. As an example one MS/MS spectrum collected from the specified 13 parent ions (including the internal standard, *S*-adenosylmethionine-*d*3) and product ions arising from their fragmentation is shown in Figure 5.3. The masses of the precursor ions and the formed product ions used for detection of the metabolites of interest are shown in Chapter 2, Table 2.6. The LC-TSQ method used was not suitable for quantification of folic acid and homocysteine. Therefore, the data for these two metabolites have not been included. The data for the 10 quantifiable metabolites of interest are shown in Figure 5.4. As shown in this figure, methionine and SAH levels were higher and statistically significant in both HCA and ST samples compared to healthy dab liver while the level of choline was statistically significantly lower in ST and HCA samples compared to healthy dab liver samples. Dimethylglycine and adenosine levels were lower in ST and higher in HCA samples compared to healthy dab liver samples (statistically significant), respectively. As shown in Figure 5.5, SAM/SAH ratio was statistically significantly lower in HCA and ST samples compared to healthy dab liver samples.

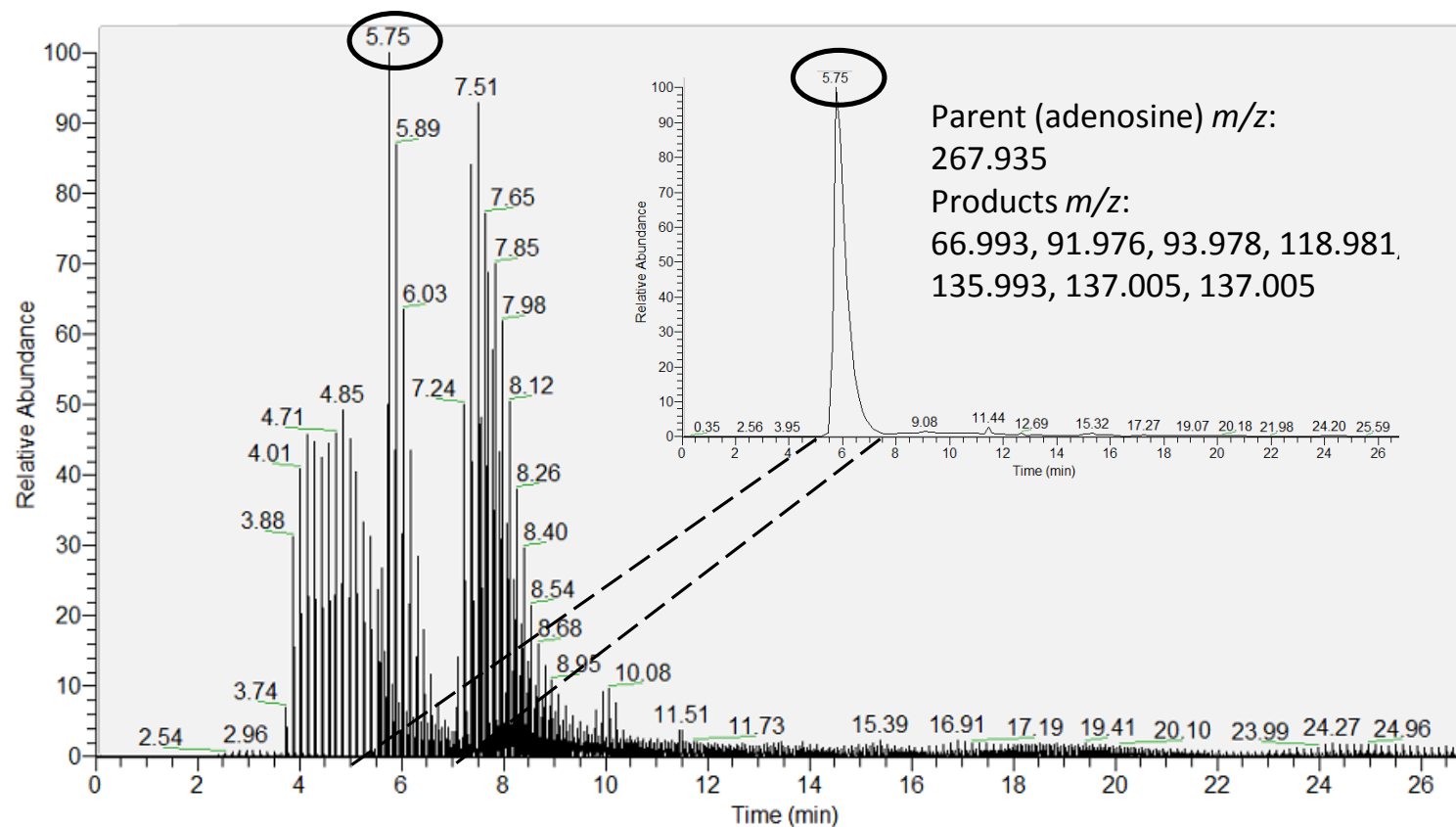
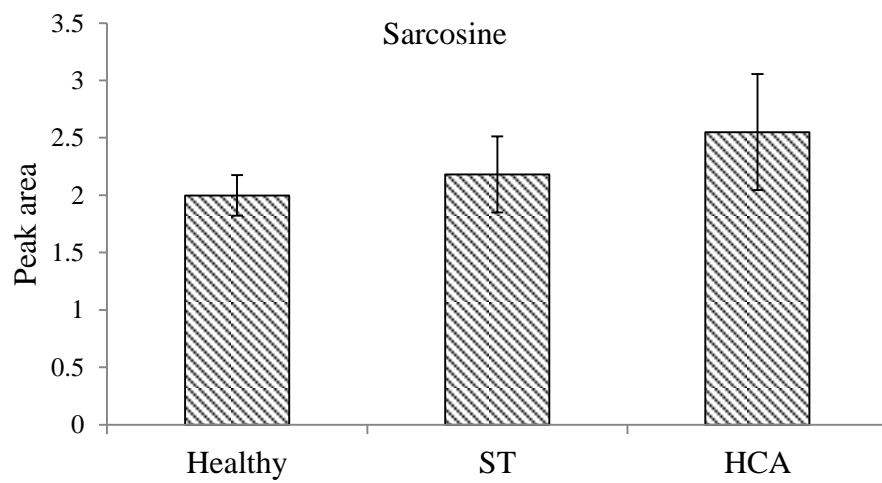
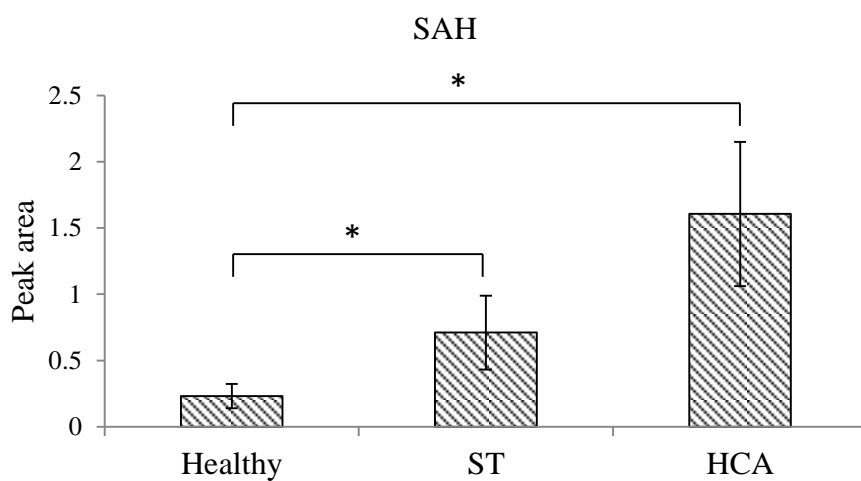
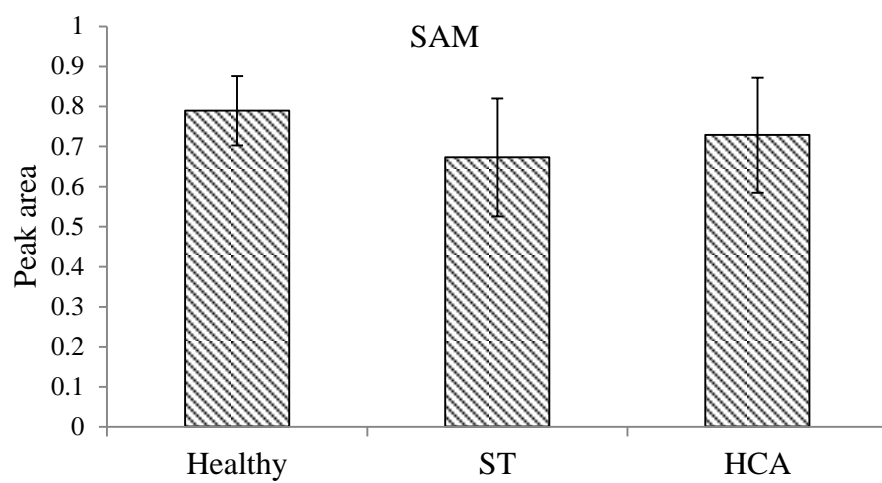
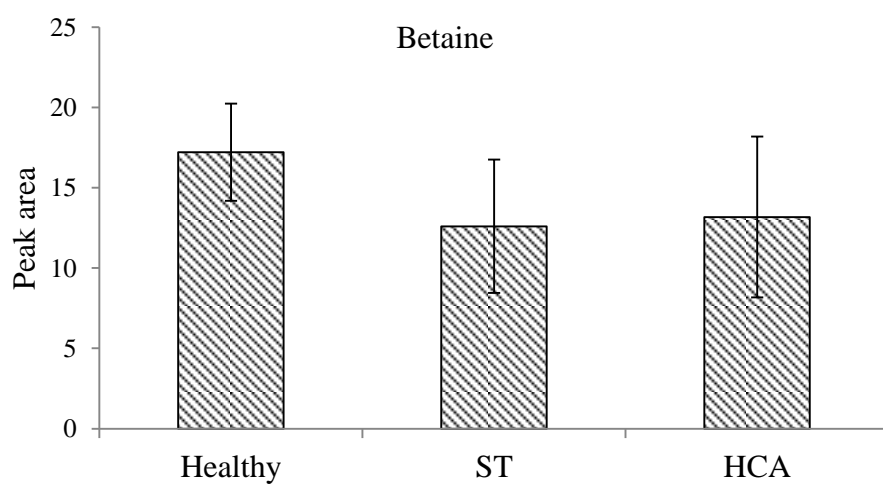
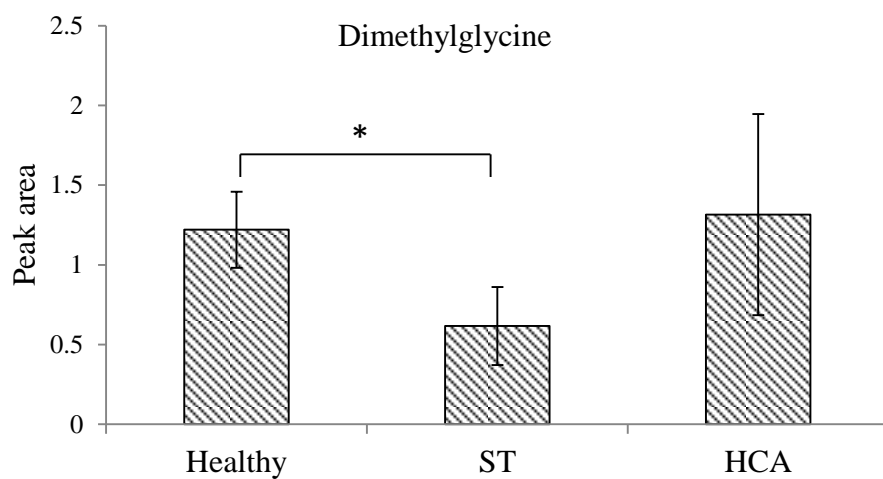
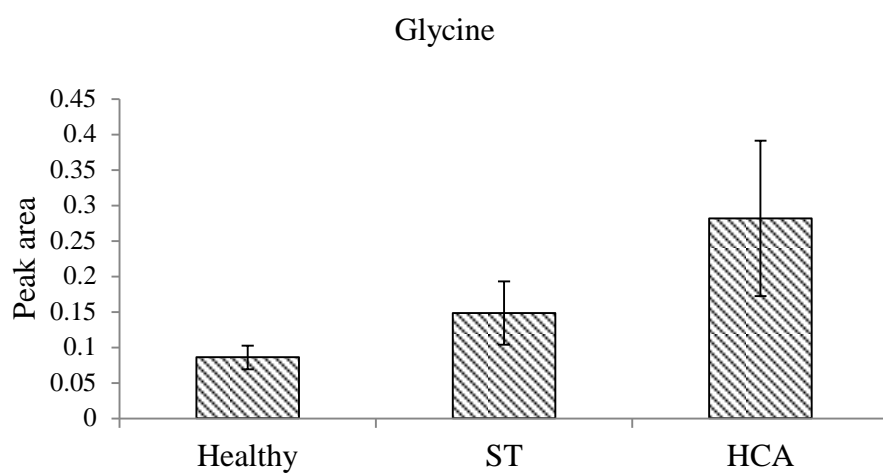
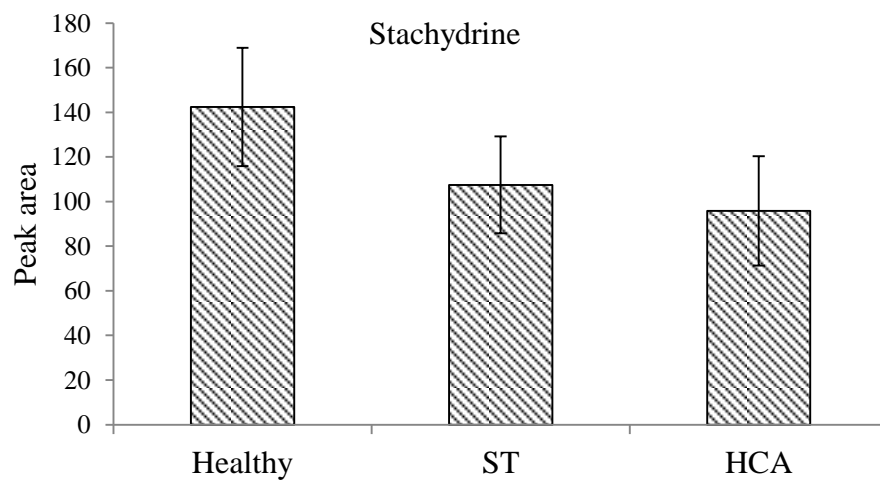
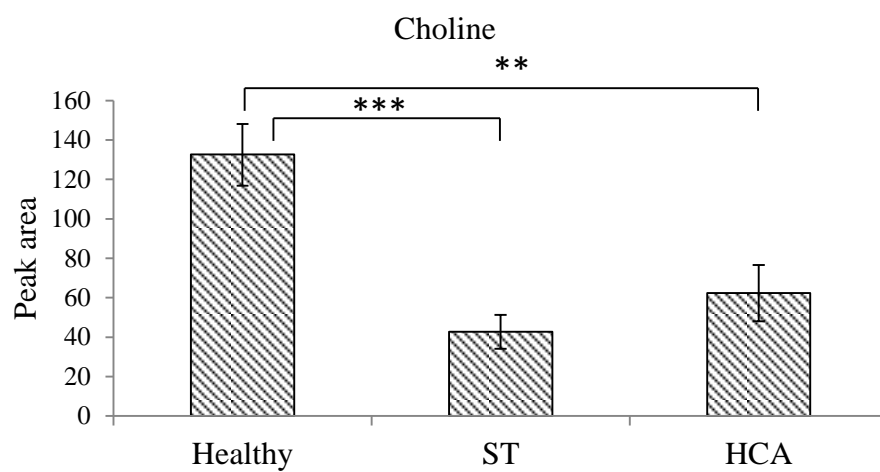
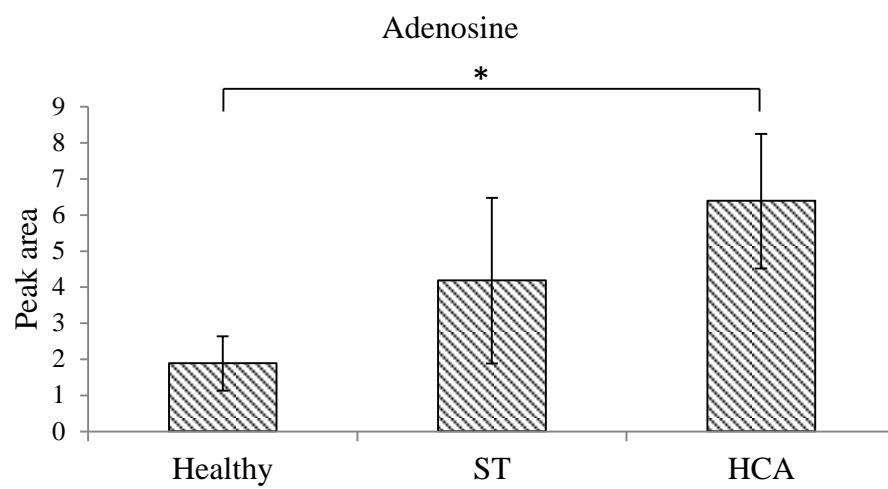


Figure 5.3. Example of a MS/MS spectrum collected from the specified 13 parent ions and product ions arising from the fragmentation of the parent ions. The collision energy was altered according to the values indicated in Chapter 2, Table 2.6, to scan (scanning time: 27 minutes) for the presence of the 13 parent compounds and their products. The vertical bars relate to the relative current produced by ions of varying m/z . Therefore, the greater the current the more abundant the ion is in the sample. For example by combining the bars related to the m/z of the adenosine (m/z : 267.935) and the seven formed ions after fragmentation of the adenosine, a peak relating to the metabolite adenosine at retention time of 5.75 was detected as shown in top right corner of the above figure. The peak area for each metabolite relates to the relative abundance of the metabolite. The m/z values for the seven adenosine product ions are shown in this figure.







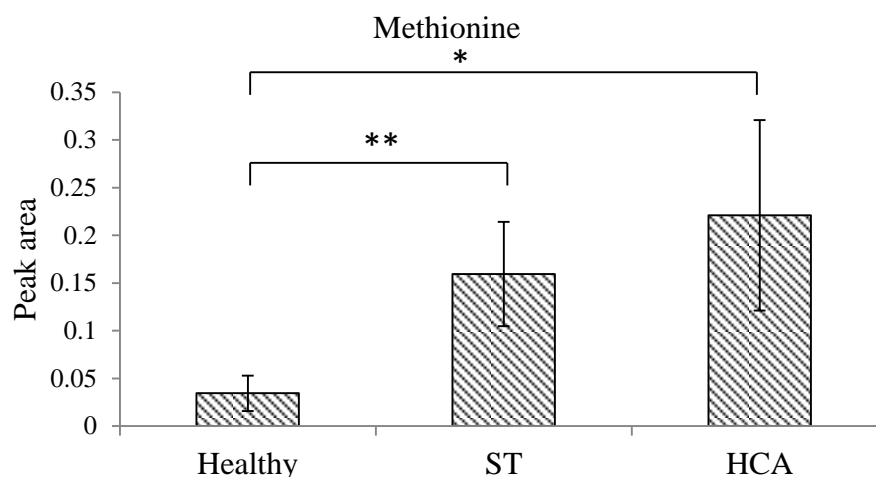


Figure 5.4. Quantification of the 10 metabolites of interest in hepatocellular adenoma (HCA), surrounding tissue (ST) and healthy dab liver samples. In these graphs, the relative abundance of each metabolite is shown, having been measured as peak area. The quantification of the peak area for each metabolite relates to 4.5 mg of dab liver sample. Each sample was spiked with the internal standard, *S*-adenosyl-*L*-methionine-*d*3 (0.016 mmol/ ml). The peak area for each metabolite was normalised to the peak area for the internal standard. The name of the metabolite investigated is shown at the top of each graph. Data is presented as mean \pm SEM of 12 healthy dab liver samples, 10 ST and 10 HCA samples. The same samples used in Chapter 4 were used for the experiments conducted in Chapter 5. Samples were run in duplicate for technical replication. Negative controls were used to monitor the quality of the extraction method. A mixture of acetonitrile: water was injected in-between each biological replicate to prevent sample carryover. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$. Depending on the normal distribution of the data and homogeneity of the variance as determined by Shapiro-Wilk test and Levenes' test, respectively, non-parametric test (Mann-Whitney U test) or parametric test (2-tailed Student's *t*-test) were used for determining the *P*-value.

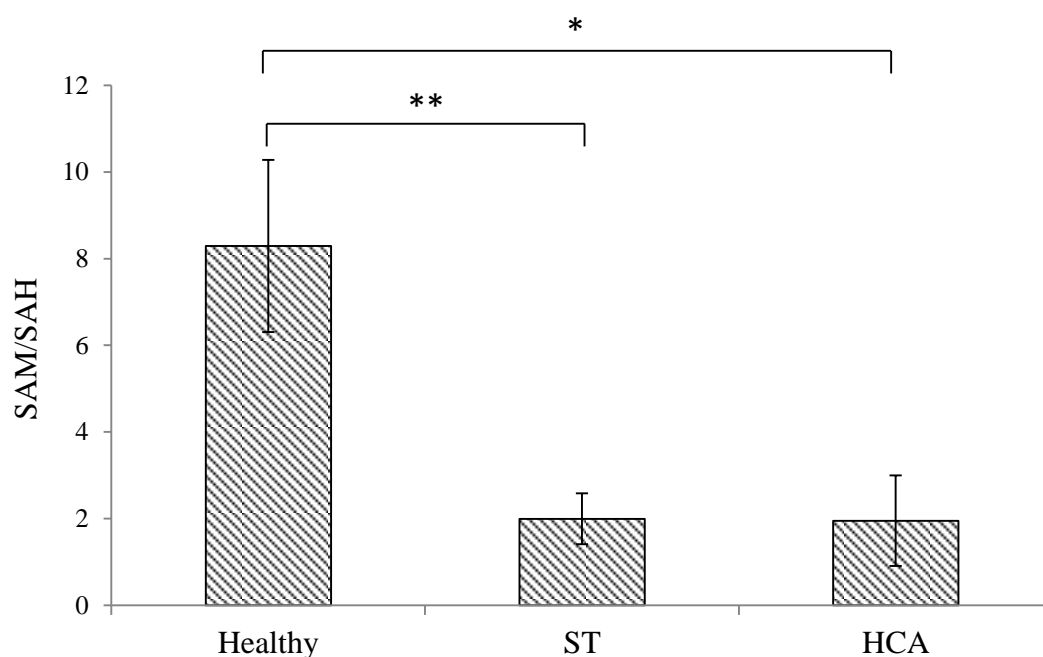


Figure 5.5. Quantification of SAM/SAH ratio in HCA, ST and healthy dab liver samples. Data are presented as mean \pm SEM of 12 healthy dab liver samples, 10 HCA and 10 ST samples. ** $P < 0.01$, * $P < 0.05$. P -value was determined using 2-tailed Student's t -test.

5.3.2. Principal components analysis

The normalised data for all the 10 metabolites investigated from the three categories of interest (HCA, ST and healthy dab liver) were subjected to principal component analysis (PCA) scores plot (Figure 5.6).

In the PCA plot, healthy dab liver samples were separated from the HCA and ST samples based on disease along the PC2 axis. PC1 and PC2 accounted for 65.8% and 32.52% of the variance, respectively. In this PCA plot, HCA and ST were clustered together. This demonstrated that the metabolic profiles of the HCA and ST samples for the investigated metabolites are similar. This result was in agreement with the data demonstrated in Figure 5.4. As shown in Figure 5.4, although some metabolites had increasing or decreasing trends in HCA samples compared to ST samples, statistically significant differences were not observed between these samples and this was reflected in the results of the PCA plot (Figure 5.6).

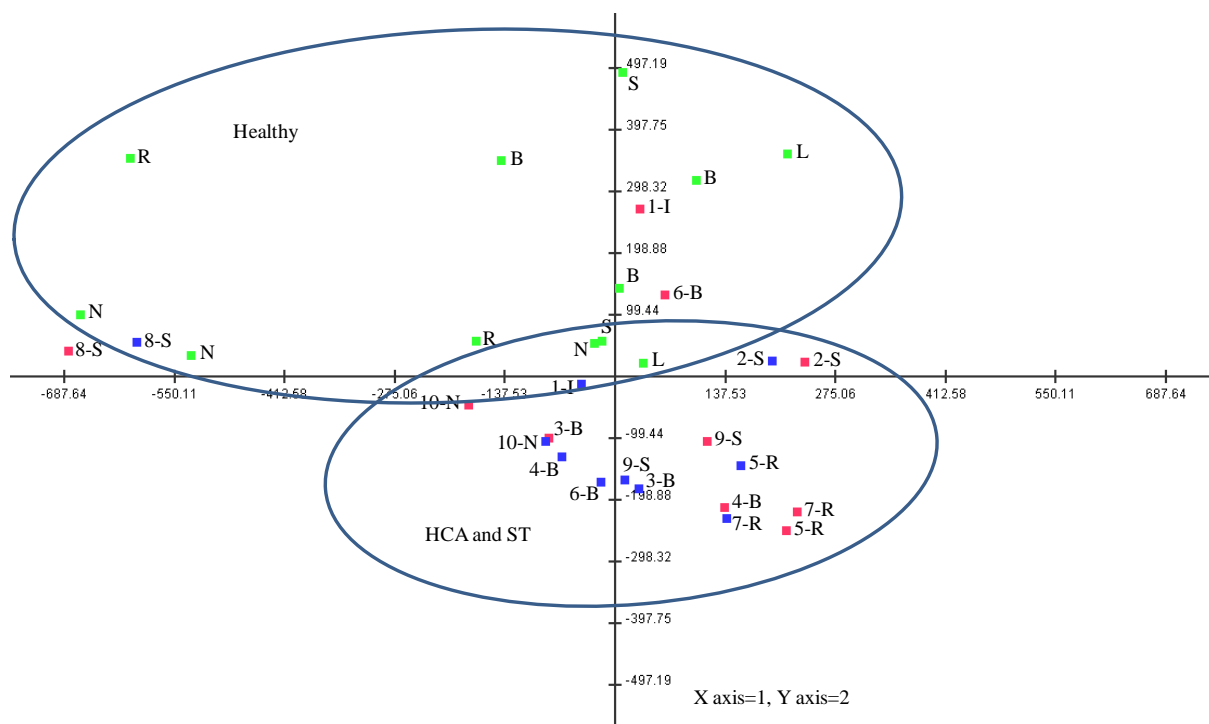


Figure 5.6. Principal components analysis (PCA) scores plot for the 10 metabolites that were investigated in the HCA, ST and healthy dab liver samples using LC-TSQ. Dab healthy liver samples (green) were separated based on disease from the HCA (red) and the surrounding tissue (blue) samples along the PC2 axis. Numbers 1-10 represent 10 individual fish. Matching numbers are ST and HCA samples from the same fish. Sampling sites for each fish is also shown. N: North Cardigan Bay, S: South Cardigan Bay, I: Inner Cardigan Bay, B: St Bee's Head, R; Red Wharf Bay, L: Lundy.

5.3.3. Change in the expression and DNA methylation levels of the genes involved in the one-carbon cycle

The expression levels of 23 probes related to one-carbon cycle for the three tissue categories of interest, represented on the dab 15K microarray (described in Chapter 4), were re-analysed. These included: *S-adenosylhomocysteine hydrolase* (probe ID: 18T4428314), *methionine adenosyltransferase* (probe ID: 6140736), *methionine adenosyltransferase 1* (probe ID: 6606586), *methionine adenosyltransferase 2* (probe ID: 6625279), *DNA methyltransferase 1* (probe IDs: 6355974 and 18T3383577), *DNA methyltransferase 3a* (probe IDs: 6396450 and 19N4016975), *DNA methyltransferase 3b* (probe IDs: 6505586, 6623925 and 6185506), *DNA methyltransferase 4, 5, 7 and 12* (probe ID: 6102058, probe ID: 6613415, probe ID: 6521192, probe ID: 6516262), *S-adenosylmethionine dependent methyltransferase* (probe ID: 6245510), *betaine-homocysteine methyltransferase* (probe IDs: 18T3874626, 19N3859050, 19N4347827, 18T4139617, 6593006), *methylenetetrahydrofolate reductase* (probe ID: 6254480), *methionine synthase* (probe ID: 18T3885734). As comparisons were restricted to limited number of probes (23 probes), a *p*-value cut-off of less than 0.05 and a fold change >1.5 with no FDR correction were applied on the normalised gene expression data achieved in Chapter 4, section 4.3.2.3. Table 5.1 demonstrates the genes that their expressions are significantly changing between the three investigated tissue categories. The fold change differences in the transcription levels of all the 23 probes are presented in Appendix, File 16.

Furthermore, probes related to the one carbon cycle with DNA methylation alterations of greater than 1.5 fold between HCA and ST as determined by MeDIP-HTS in Chapter 4, are represented in Table 5.2.

As demonstrated in Table 5.1, both over- and under-expression of members of DNMT family were observed in HCA samples compared to ST and healthy dab liver samples. While the expression of the *DNA methyltransferase 3a* gene and one of the probes matching to *DNA methyltransferase 3b* gene were statistically significantly lower in ST compared to healthy dab liver with decreasing trends in HCA samples compared to healthy dab liver samples, the expression of the hypomethylated *DNA methyltransferase 4* was higher and statistically significant in HCA samples compared to ST samples. The *S-adenosylhomocysteine hydrolase* and *betaine-homocysteine methyltransferase* genes were statistically significantly over- and under-expressed, respectively in HCA compared to both ST and healthy dab liver tissues. The

hypomethylated *methionine adenosyltransferase* gene in HCA sample was statistically significantly over-expressed in HCA compared to ST samples with increasing trends in HCA compared to healthy dab liver samples. One of the probes representing the *betaine-homocysteine methyltransferase* gene was lower expressed (statistically significant) in HCA samples compared to ST and healthy dab liver samples.

Category of comparison	Gene	Symbol	Fold change
HCA/H	S-adenosylmethionine hydrolase	SAHH	+2.41**
	Betaine-homocysteine methyltransferase	BMHT	-2.63***
	Methylenetetrahydrofolate reductase	MTHFR	+1.82***
ST/H	DNA methyltransferase 3b	DNMT3b	-2.21*
	DNA methyltransferase 3a	DNMT3a	-1.61*
HCA/ST	methionine adenosyltransferase	MAT	+1.78*
	Betaine-homocysteine methyltransferase	BMHT	-1.77*
	DNA methyltransferase 4	DNMT4	+1.34*
	S-adenosylmethionine hydrolase	SAHH	+3.25*

Table 5.1. Genes related to one-carbon pathway statistically significantly altered in expression between in the three categorise investigated (HCA, ST and H). *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$. Depending on the normal distribution of the data and homogeneity of the variance as determined by Shapiro-Wilk test and Levenes' test, respectively, non-parametric test (Mann-Whitney U test) or parametric test (2-tailed Student's t -test) were used for determining the P -value.

Gene	Symbol	Fold enrichment (HCA/ST)
SAM dependent methyltransferase	MT	-1.75
DNA methyltransferase 1	DNMT1	+1.50
DNA methyltransferase 3b	DNMT3b	-1.67
methionine adenosyltransferase	MAT	-1.54
DNA methyltransferase 4	DNMT4	-2.56
DNA methyltransferase 12	DNMT12	-1.72

Table 5.2. Genes with altered DNA methylation (> 1.5-fold enrichment) involved in the one-carbon cycle as quantified by *de novo* high throughput sequencing of 5-methylcytosine immunoprecipitated DNA.

5.4. Discussion

Change in the levels of metabolites is demonstrative of the physiological, developmental and pathological status of a cell. Metabolomic studies in tumours have provided valuable information for tumour characterisation, drug therapy and diagnosis (Griffin and Shockor, 2004). Furthermore, metabolites and nutrients that provide the primary sources of methyl group and by producing the final methyl donor, *S*-adenosylmethionine, contribute significantly to the regulation of DNA methylation (Ulrey et al., 2005). Therefore, not only disruption of the activities of epigenetic modifying enzymes, such as DNA methyltransferases, can have an adverse effect on health, but also changes in the levels of the metabolites required for the action of these enzymes can result in alterations to the epigenome (Turner, 2009). Therefore, in this chapter changes in the levels of a number of key metabolites in the one-carbon cycle were investigated. This study provided a snapshot of the relative levels of the metabolites of interest in the three investigated categories of tissue at a given time point. Therefore, although it does not directly provide information on the flux, nevertheless it provides valuable information regarding the changes to the one-carbon pathway. To achieve a better characterisation of the one-carbon pathway in dab HCA samples, the information on the changes in the levels of metabolites was combined with the changes observed in the DNA methylation and expression of the genes involved in this pathway. The overall changes observed in the one-carbon cycle are summarised and combined in the Figure 5.7.

5.4.1. *S*-adenosylmethionine (SAM), *S*-adenosylhomocysteine (SAH) and DNA methyltransferases (DNMTs)

SAM derived from the one-carbon pathway is the final methyl donor used in a range of transmethylation reactions, including methylation of DNA, RNA, phospholipids, proteins, histones and neurotransmitters (Krijt et al., 2009; Mischoulon and Fava, 2002). Transfer of the methyl group from SAM to cytosine, catalysed by DNMTs, results in the production of SAH (Kumar et al., 2008). In our experiment, the level of SAM did not differ between the three investigated categories of tissue. However, one of the key findings was the detection of a statistically significant higher level of SAH in “apparently healthy” non-cancerous surrounding tissue (ST) and an even higher value in HCA samples both compared to healthy

dab liver samples. SAH levels were 6.9 fold higher (statistically significant) in HCA compared to healthy dab liver and 3.1 fold higher (statistically significant) in ST compared to healthy dab liver. SAH has a high binding affinity for the catalytic regions of the DNMTs (Melnik et al., 2000; James et al., 2002). Therefore, it is a potent competitive inhibitor of DNMTs, which leads to DNA hypomethylation and alterations in the expression of specific genes (Krijt et al., 2009; Kumar et al., 2008; Melnik et al., 2000; James et al., 2002). This is in agreement with our findings, as HCA and ST samples were significantly hypomethylated compared to healthy dab liver samples. Furthermore, the expression levels of several genes were altered in HCA samples compared to ST and completely healthy dab liver samples as discussed in the previous chapter.

The level of SAM did not differ between the three tissue categories of interest. However, as SAH levels were significantly higher in HCA and ST samples compared to healthy dab liver tissues, this resulted in statistically significant 4.24 and 4.15 fold reductions in the SAM/SAH ratio in HCA compared to healthy dab liver samples and ST compared to healthy dab liver samples, respectively. This ratio, which sometimes is referred to as the “methylation index” (Krijt et al., 2009), has been used for the prediction of the cellular DNA methylation capacity. A decrease in this ratio has been linked with reduction in cellular DNA methylation capacity (Caudill et al., 2001; Melnik et al., 2000). However, several studies have demonstrated that a decrease in SAM/SAH ratio is only predictive of reduced DNA methylation capacity if it is caused by an increase in SAH concentration. Therefore, an increase in SAH level, independent from increase or decrease in the level of SAM, will result in reduced cellular DNA methylation capacity and global DNA hypomethylation (Caudill et al., 2001; Yi et al., 2000). For example, in human lymphocytes it has been shown that 2-fold elevation of SAH is associated with 2.6 fold reduction in the global DNA methylation levels of lymphocytes in the absence of change in the levels of SAM (Yi et al., 2000). While in Min mice, a positive correlation was observed between increase in the levels of both cellular SAH and SAM, decrease in global DNA methylation levels and induction of multiple intestinal neoplasia (Sibani et al., 2002). This indicates that most likely SAH, independent of SAM, is the key metabolite that regulates the activity of DNMTs (Ulrey et al., 2005). It has been suggested that SAM can only negatively influence the activity of DNMTs when its level falls significantly below the lower limit of its physiological range (Yi et al., 2000). Therefore,

changes in SAH levels can be a better marker of the cellular DNA methylation capacity than the methylation index (Caudill et al., 2001).

Alterations in the activities of DNMTs have a significant impact on DNA methylation and tumourigenesis. Changes in the activities of DNMTs can be a consequence of the interference of different factors with the function of DNMTs, such as an increase in SAH levels and inhibition of DNMTs (Krijt et al., 2009; Kumar et al., 2008; Melnyk et al., 2000), interaction of metals such as cadmium with DNMTs and interference by chemically-induced reactive oxygen species with the activity of DNMTs (Baccarelli and Bollati, 2009). However, changes in the expression levels of DNMTs, caused by change in their DNA methylation levels, have also been implicated in different stages of cancers in humans (Girault et al., 2003; Sun et al., 1997; Laird and Jaenisch, 1994; Momparler and Bovenzi, 2000; Robertson et al., 1999, Lorient et al., 2006). For example decrease in the expression levels of specific DNMTs, depending on the type of tissue, can lead to global DNA hypomethylation and development of tumours; while increase in the expression of other members of the DNMT family is associated with the re-methylation stage and maintenance of DNA methylation in highly proliferating tumour cells. This is discussed in detail in Chapter 4, section 4.4.5 (Lorient et al., 2006). In humans the functions of various DNMTs have been well characterised. Furthermore, DNMTs have been separated from their paralogs, tRNA methyltransferases, which use the same methyl donor, SAM (Different functions of DNMTs, such as maintenance and *de novo* DNA methylation in humans are discussed in details in Chapter 1, section 1.3.1). However, DNMTs, their biological functions and their distinction from tRNA methyltransferases have not been well characterised in fish species. For example, eight different DNMTs have been identified in zebrafish but their biological functions are far from characterised (Shimoda, 2007). As shown both in the results section and in Figure 5.7, the expression and methylation of several *dnmt* genes were altered in dab liver tumours compared to ST or healthy dab liver samples. Although the function of each *dnmt* is not known in fish, these changes suggest that the expression of these genes may be involved in increasing the levels of SAH, global DNA hypomethylation and gene specific DNA methylation changes that were detected in dab tumours.

5.4.2. The inhibitory role of homocysteine and its association with SAH

An increase in the cellular SAH level is generally recognised as a secondary effect of an increase in the levels of homocysteine (Melnik et al., 2000; James et al., 2002). As shown in Figure 5.7, SAH is converted in a reversible reaction via SAH hydrolase (SAHH) to adenosine and homocysteine. However, the equilibrium constant of SAHH favours the production of SAH. Therefore for the continuation of the DNA methylation reaction, it is necessary that the homocysteine is removed from the site of production. This will induce further conversion of SAH to homocysteine and removal of SAH from the site of production (Ulrey et al., 2005). Therefore, increase in homocysteine levels will prevent the removal of SAH, leading to accumulation of SAH and inhibition of DNA methyltransferase reactions (Caudill et al., 2001). In this study, no data were available on the levels of homocysteine in the three investigated categories of tissues. However, higher levels (statistically significant) of adenosine and *sahh* transcript were observed in dab HCA samples compared to ST and healthy dab liver samples. Considering all the above findings, the data are suggestive of a higher level of homocysteine in HCA samples compared to healthy dab liver samples. Higher levels of homocysteine in HCA samples, as discussed, could have caused a secondary increase in SAH levels, as observed. Furthermore, increase in the rate of SAH to homocysteine conversion may compensate and promote the removal of SAH for maintaining the lower level of DNA methylation detected in dab HCA and ST cells compared to healthy dab liver. However, it has to be noted that adenosine is formed both from conversion of the SAH to homocysteine and from dephosphorylation of adenosine monophosphate (AMP). It is also involved in many different pathways, including synthesis of nucleotides such as inosine (Fox and Kelly, 1978). Therefore, the higher level of adenosine observed in HCA samples compared to healthy dab liver samples could be due to alterations of several other pathways in dab HCA tumours.

Removal of homocysteine is accomplished through two distinct pathways of remethylation and conversion to methionine (choline and folate remethylation pathways) and degradation of homocysteine to cysteine (number 7 in Figure 5.7). The latter is a permanent method of removal of homocysteine while the former will subsequently lead to regeneration of homocysteine as shown in the Figure 5.7 and discussed in section 5.4.3 (Caudill et al., 2001). Permanent degradation of homocysteine is achieved through the one-way transsulfuration

pathway. During the first step, cystathionine β -synthase (CBS) converts homocysteine to cystathionine. Cystathionine is converted to cysteine and α -ketobutyrate via cystathionine γ -lyase (CGL). The generated cysteine will further contribute to the production of glutathione (Caudill et al., 2001). In humans, a positive correlation between increased levels of homocysteine and increased transcription from *CGL* and *CBS* genes has been identified (Chen et al., 2010). This is to accelerate the removal of and inhibit the accumulation of homocysteine (Durand et al., 2001). However, our data showed that the expression levels of the *cbs* and *cgl* genes were not altered in HCA and ST samples compared to healthy dab liver samples (data not shown, raw data are available from the ArrayExpress at EMBL-EBI and can be found under accession A-MEXP-2084). This may contribute to the higher levels of adenosine, SAH and methionine observed in dab HCA samples and ST samples compared to healthy dab liver samples.

5.4.3. Alterations in the levels of primary nutrient methyl donors; choline, folate and methionine

Methionine levels were significantly higher in dab HCA and ST samples compared to healthy dab liver tissues. As mentioned above, one possibility could be insufficient degradation of and subsequent increase in the level of its precursor, homocysteine, which could lead to an increase in methionine levels in tumours. However, there are three alternative pathways that could also impact the level of methionine. These three pathways are highly influenced by diet. Diet can influence DNA methylation mainly by affecting the one-carbon cycle (Vineis et al., 2011). In humans the three main sources of methyl groups are provided through nutrient intake of methionine (~10 mmol of methyl/day), folate (~5_10 mmol of methyl/day) and choline (~30 mmoles/day). Therefore, diet has a significant role in the regulation of the one-carbon cycle and DNA methylation (Niculescu and Zeisel, 2002). Furthermore, it has been estimated that inappropriate diet accounts for more than one third of deaths caused by cancer (Doll and Peto, 1981; WCRF, 2007). Changes observed in the one-carbon cycle in dab tumours compared to healthy dab liver samples may in part, be related to differences in diet that could affect the development of tumours in dab, however there was no information on the diets of the wild fish used in this study.

As mentioned in section 5.4.2, methionine is regenerated in the liver from homocysteine by two independent pathways: 1. The betaine-homocysteine methyltransferase (BHMT) reaction 2. Folate, vitamin B12-dependent methionine synthesis. (Emmert et al., 1996; Herceg and Vaissiere, 2011; the two pathways are described in details in Chapter 1, section 1.3.3).

5.4.3.1. Methionine and choline

In the first pathway, the primary source of the methyl group is provided by dietary intake of choline. Choline in mammals is used in three different pathways. 1. Donation of a methyl group and regeneration of methionine (number 2 in Figure 5.7) 2. Synthesis of phosphatidylcholine (PC), one of the most abundant phospholipid components of the cell membrane (number 8 in Figure 5.7) 3. Synthesis of the neurotransmitter acetylcholine (Michel et al., 2006). In the human liver, 60% of the choline is utilised in the methionine regeneration pathway (DeLong et al., 2002). In dab tumours the levels of metabolites, such as dimethylglycine and choline, and the transcription level of *betaine-homocysteine methyltransferase (bmht)* gene, which converts homocysteine to methionine, were lower compared to healthy dab liver. This suggests an alteration in the choline-methionine metabolic regeneration pathway in the livers of dab containing HCA tumours. This finding agrees with previous published data demonstrating down-regulation of this pathway in dab liver tumours (Southam et al., 2008). Furthermore, choline deficiency, caused as a result of deficiency in uptake or reduced levels in diet and/or disruption of the choline metabolic pathway, has been linked with development of liver cancer in humans (Michel et al., 2006). The same effect has been reported in rodents. Several studies using diets deficient in choline (Locker et al., 1986), choline and methionine (Wilson et al., 1984; Ghoshal and Farber, 1984) or folate and choline (Sibani et al., 2002) have suggested a link between choline deficiency, DNA hypomethylation and alterations in the levels of the one-carbon metabolites such as SAM and SAH as well as development of liver tumours and multiple intestinal neoplasia in mice and rats. These studies have concluded that choline deficiency is carcinogenic in rats (Locker et al., 1986). Uptake of the essential nutrient methionine, in choline deficient animals, causes an increase in methionine and homocysteine levels (Michel et al., 2006). As mentioned in section 5.4.2 this will further result in accumulation of SAH, DNA hypomethylation and development of liver tumours. This corresponds with changes observed in the one-carbon cycle in dab tumours. Therefore, the higher level of methionine observed in the tumour bearing dab livers compared

to healthy dab liver tissues could be as a result of a decrease in choline levels and/or direct increase in methionine intake and increase in homocysteine levels. Furthermore, the level of stachydrine (proline-betaine) was investigated in the three tissue categories of interest. Stachydrine, synthesised in plants (Heinzmann et al., 2010), has a negative impact on the betaine-methionine regeneration pathway (Lever and Slow, 2010). This compound reduces the amount of available betaine for the betaine-methionine regeneration pathway by increasing excretion of betaine from the kidneys (Lever and Slow, 2010; Heinzmann et al., 2010). However, the level of this compound was not significantly altered in dab tumours. This suggests that decrease in the available amount of choline is the main cause of decrease in the level of dimethylglycine and down-regulation of betaine-methionine regeneration pathway rather than increased excretion of betaine.

5.4.3.2. Folate

A third possibility is change in the folate-methionine regeneration pathway (number 3 in Figure 5.7). Folate is essential in regulation of DNA repair, DNA synthesis and methylation reactions. As a result, folate deficiency is generally associated with reduced production of SAM, global DNA hypomethylation, gene specific hypermethylation and increased risk of several cancers, such as breast, brain, ovary, colorectum and pancreas in humans and liver tumours in rodents (Duthie et al., 2011; Caudill et al., 2001). However, the extent to which folate deficiency affects DNA methylation is less consistent between different studies. It is thought that the damaging effect of folate deficiency depends on a combination of factors such as the type of tissue and the treatment regime. Furthermore, folate deficiency is mainly associated with global DNA hypomethylation when it is severely limited in the diet (Duthie et al., 2011). In this study the HPLC running conditions could not be optimised for measurement of folate. However, as shown in Figure 5.6, the *Methylenetetrahydrofolate reductase (mthfr)* gene was over transcribed in HCA samples compared to healthy dab liver samples. Therefore, this may suggest up-regulation of the folate-methionine regeneration pathway in the livers of dab containing tumours.

One possible outcome of up-regulation of folate-methionine regeneration pathway is to compensate for the alterations occurring in choline-methionine regeneration and homocysteine degradation pathways in dab tumours. Exogenous restriction of choline or

methionine causes an increase in synthesis of 5-methyl THF (Durand et al., 2001). The link between 5-methyl THF, SAM and the activity of glycine N-methyltransferase (GNMT) is well established. As shown in Figure 5.7, 5-methyl THF, synthesised by MTHFR from N5, N10-methylene THF has an inhibitory effect on GNMT. Inhibition of GNMT reduces conversion of SAM to SAH and increases the availability of SAM. This will lead to maintenance of methylation reactions (Ulrey et al., 2005; Durand et al., 2001). This is in agreement with: 1. Observed increase in transcription level of *methfr* gene in HCA samples compared to ST. 2. Detection of a higher level of glycine (trend, not statistically significant) in dab tumour samples compared to ST and healthy dab liver samples. 3. Same global DNA methylation levels in ST and HCA samples.

5.4.4. Production of SAM and the competition between DNMTs, GAMT and PEMT for consumption of SAM

The conversion of methionine to SAM is catalysed by methionine adenosyltransferase (MAT). The subunits for the three MAT isoenzymes found in mammals are encoded by two different *MAT* genes, *MAT1A* and *MAT2A* (Lu, 2000). However, it is not known how many *mat* genes exist in dab. In this study, the level of the precursor compound, methionine, and the expression level of one of the hypomethylated probes corresponding to the *mat* gene were higher in HCA samples compared to ST samples. This implies a possible up-regulation in the production of SAM. However the levels of SAM did not differ between the three investigated tissue categories. This suggests a simultaneous and equal increase in the demand for and usage of SAM. Considering that: 1. SAH levels were higher in dab tumour samples compared to healthy dab liver samples, which inhibit DNA methylation reactions. 2. *methfr* gene was over-expressed in dab tumours compared to healthy dab liver samples, which also inhibits DNA methylation reactions 3. Global DNA methylation levels were lower in tumour bearing dab livers compared to healthy dab liver samples; without doubt, the rate of DNA methylation is reduced in these tumours. Despite the evidence for reduced DNA methylation, levels of SAM did not differ between the different tissue types. On the basis of relatively high levels of methionine and *mat* expression in tumour and ST samples, one has to question the fate of SAM that is expected to be produced. As discussed in detail below, other methylation pathways exist (numbers 4 and 5 in Figure 5.7) that utilise SAM and hence alternative requirements may have been elevated. The key conclusion is that a deficiency of SAM is not

evident and therefore the choline deficiency (described in section 5.4.3) is not influencing DNA methylation through causing SAM deficiency. Choline deficiency potentially could be contributing to DNA hypomethylation via SAH elevation, with it being upstream of homocysteine production.

There are two other main methyltransferases in the liver that use SAM as their methyl donor. These include guanidinoacetate methyltransferase (GAMT) and phosphatidylethanolamine *N*-methyltransferase (PEMT) (Brosnan et al., 2004). In addition to the role of choline in regeneration of methionine, as discussed in section 5.4.3, it is used in the cytidine diphosphate (CDP)-choline pathway (number 8 in Figure 5.7). CDP-choline pathway is the main metabolic pathway in the liver for the synthesis of the membrane phospholipid, phosphatidylcholine (PC) (Michel et al., 2006). However, during choline deficiency, PC synthesis is catalysed by PEMT in a three step reaction (number 4 in Figure 5.7). In this reaction three SAM molecules are used as methyl donors for the production of one PC molecule from phosphatidylethanolamine (PE) (Michel et al., 2006; Brosnan et al., 2004). This leads to an increase in homocysteine levels (Michel et al., 2006). DNA methyltransferases and other methyltransferases such as GAMT and PEMT are in competition for consumption of SAM (Ulrey et al., 2005). Therefore, as discussed in section 5.4.2, increased homocysteine levels as a result of increase in PEMT activity could subsequently lead to increase in SAH and inhibition of DNA methylation. As choline and global DNA methylation levels were significantly lower and SAH level was significantly higher in dab HCA samples compared to healthy dab liver samples, this suggests that the PE-PC pathway could be increased in dab tumours. Increase in PC synthesis reflects the higher demand for the structural phospholipid, PC, by proliferating tumour cells (Southam et al., 2008). Furthermore, the synthesis of creatine from guanidine acid, catalysed by GAMT in the liver, consumes the majority of SAM (Ulrey et al., 2005; Stead et al., 2001; number 5 in Figure 5.7). In addition, increase in the activity of GAMT causes an increase in the homocysteine levels and competitive inhibition of DNMTs (Brosnan et al., 2004). Therefore, in addition to providing another mechanism for reduction in DNA methylation levels; up-regulation of these pathways provides a possible explanation for the observed increase in methionine-SAM reaction without any change to the level of SAM in dab tumours compared to healthy dab liver samples (Figure 5.7).

Figure 5.7. Overview of the changes observed in dab HCA, ST and healthy dab liver samples as investigated in Chapter 5. This figure demonstrates all the significant changes detected in gene expression, DNA methylation and the levels of metabolites in the one-carbon cycle between the three investigated tissue categories of interest (HCA, ST and healthy dab liver (H)).

Pathways: **1:** methionine cycle, **2:** betaine-methionine regeneration pathway, **3:** folate pathway (the one-carbon cycle is comprised of pathways 1, 2 and 3), **4:** PC synthesis using SAM as methyl donor, **5:** creatine synthesis using SAM as methyl donor, **6:** DNA methylation using SAM as methyl donor (4, 5 and 6 are three reactions catalysed by three SAM-dependent methyltransferases), **7:** permanent removal of homocysteine, **8:** cytidine diphosphate (CDP)-choline pathway.

Arrows: decrease or increase. Different colours are used for the arrows to indicate changes in metabolites, gene expression and DNA methylation as well as the compared categories of tissues (i.e. HCA vs. ST, HCA vs. H or ST vs. H).

Abbreviations: ADP: adenosine diphosphate, ATP: adenosine triphosphate, BMHT: betaine-homocysteine methyltransferase, CDP-choline: cytidine diphosphate-choline, CK: choline kinase, CMP: cytidine monophosphate, CPT: cholinephosphotransferase, CTP: cytidine triphosphate, DNMT: DNA methyltransferase, DAG: diacylglycerol, GAA: guanidinoacetate, GAMT: guanidinoacetate methyltransferase, GNMT: glycine *N*-methyltransferase, MAT: methionine adenosyltransferases, MS: methionine synthase, MTHFR: methylenetetrahydrofolate reductase, PC: phosphatidylcholine; Pcytl: phosphocholine cytidyltransferase 1, PE: phosphatidylethanolamine, PEMT: phosphatidylethanolamine *N*-methyltransferase, PPi: pyrophosphate, SAH: *S*-adenosylhomocysteine, SAHH: *S*-adenosylhomocysteine hydrolase, SAM dependent MT: *S*-adenosylmethionine dependent methyltransferase, THF: tetrahydrofolate.

5.4.4. Conclusions

In this study, it was demonstrated that levels of the components of the one-carbon cycle differed in tumour bearing dab livers compared to healthy dab liver samples. The two most apparent significant differences were detection of higher levels of SAH and lower levels of choline in dab HCA and ST compared to healthy dab liver. The former is a biomarker of reduced capacity of cellular DNA methylation (Caudill et al., 2001; Yi et al., 2000). The latter indicates either reduced cellular uptake of choline in the liver or choline deficiency in the diet. Reduction in the bioavailability of choline for the one-carbon cycle has also been linked with DNA hypomethylation and development of tumours (Michel et al., 2006; Locker et al., 1986;

Wilson et al., 1984; Ghoshal and Farber, 1984; Sibani et al., 2002). As these tumours are detected in wild fish, without doubt development of these tumours is multifactorial. Therefore, in addition to the presence of pollutants in their environment (discussed in Chapter 4), it is possible that dietary factors are important in higher tumour susceptibility of fish collected from certain sampling sites. However, further studies are required to distinguish between the changes in the one-carbon pathway that are involved in development of these tumours from the changes that are a consequence of the formed tumours. In part, this has been investigated by including apparently healthy tissue (ST) from tumour containing liver and completely healthy liver. Based on the PCA results, HCA and ST samples are metabolically similar in regards to the one-carbon metabolic pathway and both differed from healthy, non-cancerous dab liver tissues. This was reflected in the graphs shown above, as no significant differences were observed between HCA and ST samples. Detection of similar profiles between HCA samples and non-cancerous ST samples, suggest that the differences observed in the levels of metabolites between HCA and healthy dab liver samples could be implicated in the development of tumours. However, further studies using pre-neoplastic lesions are needed before any decisive conclusions on the earlier stages of tumourigenesis could be made.

Chapter 6

General discussion and future work

6.1. General discussion

Assessment of the disease status of common dab (*Limanda limanda*), captured from the waters around the UK, is used in the UK Clean Seas Environmental Monitoring Programme (CSEMP) as an indicator of biological effects of contaminants in the marine environment. Surprisingly, a high prevalence of liver tumours, with some sites exceeding more than 20%, has been reported in this fish species (described in Chapter 1, section 1.4.1.2). However, the underlying causative factor/factors are not known.

Overlaid on the genome there are layers of heritable chemical signals and markers referred to as epigenetic markers (Bernstein et al., 2007). As described in section 1.3.2, DNA methylation, as an epigenetic mechanism in connection with other epigenetic and non-epigenetic factors, is important in control and regulation of transcription. Therefore, disruption of DNA methylation has been associated with development of a range of disorders in humans (Egger et al., 2004), including cancers (De Smet and Loriot, 2010; Feinberg et al., 2006). It is becoming more apparent that aberrant DNA methylation is a common and early event during tumourigenesis and that epigenetic abnormalities increase as the tumours progress (Lempiainen et al., 2011; discussed in Chapter 1, section 1.3.5.3). In addition, as described in section 1.3.5.2, epigenetic factors are important in linking the signals received from the environment to phenotypic responses; hence, providing a mechanism for adaptation to varying environmental conditions that the genome alone cannot provide.

As discussed in Chapter 1, section 1.3.1, Table 1.1, there is very little information on epigenetic changes in aquatic species as a response to exposure to model environmental pollutants. Epigenetic is an emerging field in aquatic biology and several aspects have been neglected thus far, including the role of epigenetic alterations in diseases of aquatic species. To our knowledge, no previous DNA methylation studies of tumours dissected from wild fish have been reported. Therefore, the work presented in this thesis investigated the possibility of alterations in DNA methylation profiles of tumours in fish and provided a comprehensive characterisation of liver tumours in dab in regards to changes in DNA methylation, transcription and the one-carbon methylation pathway, providing a hypothesis for formation of these tumours.

6.2. Main findings on global DNA methylation

In Chapter 3, it was shown that, in agreement with the literature (Aniagu et al., 2008; Bird, 1980), overall DNA methylation levels are two fold higher in the livers of fish (8%) compared to mammals (4%), including human liver. It is thought that body temperature is positively correlated with deamination rate of 5-methylcytosine and 5mCpG:TpG transition and therefore negatively correlated with overall DNA methylation levels. As a result, TpG and CpA dinucleotides are over-represented and 5-mCpG dinucleotide is under-represented in the genome of homeothermic species compared to the genome of ectothermic species (Varriable and Bernardi, 2006a; Varriable and Bernardi, 2006b; Jabbari and Bernardi, 2004; Bird, 1980). In Chapter 4, it was demonstrated that overall DNA methylation levels were significantly lower (~1.8 fold) in both ST and HCA samples compared to healthy dab liver. This is similar to the detection of severe DNA hypomethylation in human cancers (Lempiainen et al., 2011). The mechanism of global DNA hypomethylation and gene specific hypermethylation in human cancers is explained in a model referred to as demethylation-remethylation model (The impact of global DNA hypomethylation and its role in tumourigenesis is described in Chapter 1, section 1.3.5.1.1). As described in Chapter 1 (section 1.3.5.1.1 and Figure 1.12) and Chapter 4 (section 4.4.5) in this model, tumour cells undergo an early transient stage of genome wide de-methylation followed by selective re-methylation and maintenance of DNA methylation. This model also takes into account the changes observed in the expression of DNMTs in different stages of tumourigenesis (Loriot et al., 2006; De Smet and Loriot, 2010). Furthermore, the demethylation-remethylation model describes how it is possible that both hypo- and hyper-methylated gene-specific regions coexist in human tumours (Loriot et al., 2006). Although the DNA extracted from ST and HCA samples was globally hypomethylated, similar to cancers in humans, gene-specific hypermethylated regions were detected in dab tumours compared to ST and healthy dab liver. Furthermore, the expression and methylation of *dnmt* genes, as described in section 4.4.5, were altered in dab tumours. Therefore, it is proposed that possibly the same mechanism of demethylation-remethylation occurs in dab tumours as in mammalian tumours, which may explain the co-occurrence of both hypo- and hyper-methylated regions in dab tumours.

6.3. Zebrafish and dab DNA methylation studies

Dab is a non-model, un-sequenced flatfish. Therefore, any gene-specific DNA methylation studies at the gene level are challenging. In contrast, the zebrafish genome is sequenced and due to significant histopathological similarities between zebrafish tumours and corresponding human tumours (Lam and Gong, 2006), it is commonly used as a model for studying human tumourigenesis (Feitsma and Cuppen, 2008; Lawrence et al., 2009). However, there was no information available on the role of DNA methylation abnormalities in development of tumours in zebrafish. Therefore in Chapter 3, DNA methylation alterations were characterised in chemically induced HCC tumours in comparison with completely healthy zebrafish livers. The data gathered from this study not only provided the first evidence of alterations of DNA methylation in tumours of aquatic species but also indicated that DNA methylation alterations could have a significant role in the development of tumours of other fish species, such as dab. This strengthened the need for a methodology to study DNA methylation at the gene level in un-sequenced fish species. As a result in Chapter 4, MeDIP was combined for the first time with *de novo* high-throughput DNA sequencing (HTS). Gene specific studies in dab and zebrafish, as described in Chapters 3 and 4 (sections 3.4.2, 3.4.4 and 4.4.4), demonstrated that the expression and DNA methylation of genes involved in key pathways associated with cancer were altered in both species. For example, both Wnt- β -catenin and MAPK pathways were altered in dab and zebrafish tumours in regards to DNA methylation and transcription levels. Both pathways are associated with proliferation, apoptosis, differentiation and migration in cancers and are also commonly deregulated in human liver cancers (Dhillon et al., 2007; Firestein et al., 2008; Bikkavilli and Malbon, 2009; Lam et al., 2006). This finding is not unexpected as both pathways, like other parameters related to protein synthesis and energy production, are features that give tumour cells a selective advantage for clonal expansion and are required for high levels of proliferation. Furthermore, comparisons of DNA methylation profiles of human healthy liver and HCC with zebrafish healthy liver and HCC demonstrated that although at the individual gene level zebrafish and humans are different in regards to DNA methylation profiles, similar families of genes are altered in terms of DNA methylation in both species.

Neither of the studies identified a significant overall inverse correlation between gene expression and DNA methylation. This is in agreement with other genome-wide methylome

and transcriptome studies (Lempiainen et al., 2011; Klug et al., 2010). As discussed in Chapter 1 (section 1.3.2.3) and Chapter 3 (section 3.4.4) this is not un-expected as the relationship between gene expression levels and DNA methylation is extremely complex and site sensitive (Van Vlodrop, 2011; Jones and Takai, 2001). As discussed in Chapter 1 (section 1.3.2.3) several factors, such as identification of core CpG sites and the levels of methylation required for inactivation of transcription for each gene, are key in interpretation of the exact effects of DNA methylation upon transcription. Furthermore, the location of DNA methylation in respect to the transcription start site, the presence of core CpG sites within alternative promoters at intra- and intergenic CpG dinucleotides, the regulation of transcription of several genes by one promoter and the location of several core CpG sites outside the traditionally defined CGIs, are amongst many factors adding to the complexity of the relationship between gene expression and DNA methylation (Van Vlodrop, 2011; Siegfried and Simon, 2010; Jones, 1999; Jones and Takai, 2001).

6.4. Hypotheses for formation of tumours in dab liver

As discussed in Chapter 4 (section 4.4.6), several lines of evidence including; 1. Higher prevalence of tumours in dab sampled from sites (i.e. Cardigan Bay) with higher levels of pollutants such as EDCs (e.g. PCBs and PBDEs) and metals (e.g. Cd), 2. Increased expression of *vtg*, the biomarker of exposure to estrogen mimicking compounds, in tumour-bearing dab livers compared to healthy dab livers and 3. Detection of intersex by other research groups in these fish populations (Scott et al., 2007), indicated that chronic exposure to estrogenic pollutants may contribute to the development of liver tumours. Several studies have identified a link between exposure to chemicals such as EDCs (PCB, EE, E2) and heavy metals (e.g. As, Cd), changes in DNA methylation (Vandeghechuchte and Janssen, 2011) and tumourigenesis in rodents and humans (Arita and Costa, 2009; Li et al., 2003; Lo and Sukumar, 2008; Skinner et al., 2010; Baccarelli and Bollati, 2009). For example, early postnatal exposure of rat neonates to environmentally relevant concentrations of the endocrine disrupters bisphenol A or 17 β -estradiol, induced DNA methylation changes in a range of genes, including genes involved in cell signalling, as early as postnatal day 10. Furthermore, these changes increased the risk of adult onset of precancerous lesions and hormonal carcinogenesis in the prostate of the exposed rats (Ho et al., 2006). In our study, DNA methylation and expression of several genes in both the non-genomic and genomic estrogenic pathways, possibly stimulated by estrogenic

compounds, were altered in dab tumours. These genes are involved in biological functions such as proliferation and growth, differentiation, cell cycle, apoptosis and oncogenesis (Bjornstrom and Sjoberg, 2005; Obrero et al., 2002; Safe and Kim, 2008). It is plausible that chronic exposure to EDCs, amongst other possible chemical contaminants, resulted in an initial alteration in DNA methylation. This is supported by detection of global hypomethylation in ST (Chapter 4, section 4.3.1) and significant alterations to the one-carbon cycle in ST, including a statistically significant higher level of SAH in dab tumours compared to healthy dab liver (Chapter 5, section 5.3.1). This suggests that the livers of these fish were epigenetically primed and more susceptible to development of tumours (further discussed in section 6.5). Additional epigenetic and genetic changes, for example in the non-genomic and genomic estrogenic pathway, may lead to oncogenesis in the liver of flatfish dab (as discussed in Chapter 4, section 4.4.6). In addition, the high occurrence of tumours in the livers of flatfish dab and the lack of detection of mutations in several commonly mutated genes in human cancers in the tumours of flatfish species (Rotchell et al., 2001a; Franklin et al., 2000), suggests that alterations in DNA methylation might have a key role in development of these tumours.

These fish were captured from their natural habitat. Therefore, it is most likely that formation of liver tumours is multifactorial and elements other than exposure to pollutants are involved in the formation of these tumours. Indeed, the data presented in Chapter 5 as well as demonstrating a significant role for DNA methylation in development of liver tumours, showed a possible influence of diet on development of these tumours. The level of choline, for example, was statistically significantly lower in dab tumours compared to healthy dab liver. As discussed in Chapter 5 (section 5.4.3), this could be either as a result of reduced cellular uptake or deficiency in diet. Several studies have demonstrated a link between diets deficient in choline or other primary methyl donors, DNA hypomethylation and the development of tumours in rodents and humans (Michel et al., 2006; Locker et al., 1986; Wilson et al., 1984; Ghoshal and Farber, 1984; Sibani et al., 2002). However, no information was available on the diet of these fish (eg. stomach contents). Therefore, further research is required before any conclusions could be made on the effect of diet on liver tumour susceptibility in dab.

One important question is, “are the changes observed in the DNA methylation and expression of the genes involved in the genomic and non-genomic estrogenic pathway in dab tumours randomly caused or has this pathway been specifically affected due to exposure to EDCs, i.e. does each category of chemical compounds induce distinct DNA methylation and subsequent transcriptional changes upon exposure?” Phenobarbital (PB) is a non-genotoxic compound that can cause tumours through the constitutive androstane receptor (CAR) pathway in the liver of treated rodents. Lempiainen et al., (2011) demonstrated that treatment with PB (in the absence of a mutagenic compound) causes non-random and tissue-specific changes in DNA methylation and transcription. While Cyp2b10, one of the main genes affected through the CAR pathway, was hypomethylated and over-expressed in the liver of treated mice, it was not affected in non-tumour bearing kidney. Furthermore, there was no significant overlap between DNA methylation or transcriptional changes observed in the liver and kidney of the treated mice. Several other studies have also indicated that at various stages of PB-induced tumourigenesis DNA methylation and transcription changes are non-random (in the absence or presence of mutagenic compounds) (Phillips and Goodman, 2009a; Phillips et al., 2009b; Bachman et al., 2006). Further evidence on the dependency of DNA methylation profile on the causative agent comes from a study demonstrating that HCC tumours induced by hepatitis C virus (HCV) or hepatitis B virus (HBV) have different methylation profiles (Pei et al., 2009). These studies all give rise to the idea that each compound has a specific DNA methylation “fingerprint”. This further supports our finding that DNA methylation and changes in gene expression observed in dab tumours were enriched in genes of particular pathways, for example non-genomic and genomic estrogenic pathway and pathways commonly associated with general hallmarks of cancer.

However, the initial DNA methylation changes (referred to as driver methylation) should be separated from DNA methylation changes that occur as a consequence of tumour formation (referred to as passenger methylation) (Herceg and Vaissiere, 2011). As mentioned earlier, some DNA methylation changes (e.g. changes in genes of the energy pathways) are not unexpected. These changes provide tumour cells with a selective advantage for clonal expansion. Therefore, to identify appropriate genes with DNA methylation changes as biomarkers of exposure to non-genotoxic carcinogens, it is important to separate them from passenger methylations and DNA methylation changes that are dependent on the type of tumour (described in Chapter 1, Table 1.3). Furthermore, tumours are comprised of different

cell populations (e.g. inflammatory cells, hepatic cells, etc). Differentiation and cell characterisation is highly regulated by epigenetic mechanisms as described in Chapter 1 (section 1.3.4.3). Therefore, it is possible that some of the DNA methylation changes merely reflect the differences in cell types. However, most likely these changes would be outweighed by the scale of the differences in DNA methylation between tumour and healthy cells.

6.5. The importance of ST in this study and the impact of the epigenetic progenitor model of tumourigenesis

As described in Chapter 1 (section 1.3.5.3), a significant change in the order of events leading to tumourigenesis has been proposed. This change is based on the recognition of the fact that epigenetic abnormalities, independent of changes in the genetic material, can cause a stable and sufficient change in gene expression which subsequently can lead to alteration in the phenotype (Moggs et al., 2004). Furthermore, it is becoming widely accepted that epigenetic alterations, especially the initial global DNA hypomethylation, occur at early stages of tumourigenesis even prior to detection of mutations. Evidence, including from our studies conducted in “apparently healthy” surrounding tissues as described in Chapter 1 (section 1.3.5.3) and Chapter 4 (sections 4.4.1) supports this claim. These various studies support the proposal of an alternative model of tumourigenesis known as epigenetic progenitor model of tumourigenesis suggested by Feinberg et al (2006). It was proposed in chapter 4 (sections 4.4.1 and 4.4.7) that development of dab HCA tumours appears to support the epigenetic progenitor model of tumourigenesis. This was in part based on the detection of statistically significant global DNA hypomethylation in ST compared to healthy dab liver and the detection of gene-specific differences in DNA methylation levels, in the absence of change in the overall DNA methylation levels, in the ST compared to HCA samples. As demonstrated by Ho et al (2006) in rats and described in section 6.4, early life stage exposure to epigenetic modifying chemicals such as EDCs can cause hormonal-dependent adult onset of tumours. Detection of similar lower levels of global DNA methylation in dab ST and HCA samples is also suggestive of the possibility of early life stage changes to the epigenome as a result of exposure to environmental pollutants, such as EDCs and higher susceptibility to hormonal- or age-dependent adult onset of liver tumours in this fish species.

However, one possibility that could also explain the observed global DNA hypomethylation in ST and contradict the epigenetic progenitor model of tumourigenesis, is that the detected global DNA hypomethylation is secondary to formation of tumours (i.e. the tumour has caused hypomethylation of the adjacent tissue). However, the tissues categorised as ST (used in Chapters 4 and 5), are from any histopathologically healthy section of dab liver and are not necessarily immediately adjacent to the tumour. This, although not entirely, minimises the possible secondary effects of tumours. Additionally, it was demonstrated in Chapter 5 that although HCA and ST are histopathologically different, their metabolic profiles within the one-carbon cycle are similar and both differed from healthy dab liver (while ST and completely healthy dab liver are histopathologically similar). This, as well as strengthening the theory proposed in Chapter 4 (section 4.4.7), demonstrates the importance of using both ST and completely healthy liver samples in any studies. For example, in the study conducted by Southam et al. (2008), apparently healthy tissue dissected from the tumour-containing dab liver was used as control tissue for characterising the metabolic changes during hepatic tumourigenesis in dab. However, as demonstrated in this thesis and implied by the epigenetic progenitor model of tumourigenesis, ST cannot be viewed as a healthy control tissue. Therefore, caution should be employed in the interpretation of such studies.

As mentioned in Chapter 1 (section 1.3.5.3) and Chapter 4 (section 4.4.8), the data presented in this thesis and the acceptance of the epigenetic progenitor model of tumourigenesis have several outcomes. This model will change our understanding of the biology of cancer cells. Identification of appropriate epigenetic biomarkers can lead to detection of tumours earlier than the stage currently known as benign pre-neoplastic lesion (Feinberg et al., 2006; Herceg and Vaissiere, 2011). Furthermore, early epigenetic changes in cancer cells are reversible. This provides a suitable treatment opportunity and new targets for design of antineoplastic drugs (the design of HDAC inhibitors and DNA methyltransferase inhibitors are currently an active research area) (Moggs et al., 2004; Feinberg et al., 2006). Furthermore, in the context of toxicity testing, whether in laboratory animals to assess the risk to humans or in aquatic species relevant to the natural environment, epigenetic changes are not currently a standard feature of safety assessment. Part of the reason for this is the inability to interpret the findings in relation to potential adverse outcome without a more complete knowledge of the fundamental relationships between specific changes and disease. There is also the need for identification of suitable model systems for evaluation of chemical induction of epigenetic

changes as well as suitable techniques and endpoints (Goodman et al., 2010). However, the accumulating evidence for the contribution of DNA methylation alteration in tumourigenesis and adult-onset of disease produced by numerous non-genotoxic and genotoxic carcinogens (Goodman et al., 2010; Moggs et al., 2004; Legler, 2010; Guerrero-Bosagna et al., 2005) and the established carcinogenic effect of DNA methylation changes following deficiencies in choline and other primary methyl-donors (Michel et al., 2006; Locker et al., 1986; Wilson et al., 1984; Ghoshal and Farber, 1984; Sibani et al., 2002), indicates that such mechanisms should not be ignored. Nevertheless, this does not necessarily imply a concern in relation to the aquatic environment regarding methylation changes in the absence of genetic toxicity. Unlike the assumptions made about genotoxic carcinogens, a pragmatic threshold level of exposure to non-genotoxic carcinogens or dietary deficiency may be required for a clear impact on cancer development. However, methylation of CpG islands increases the rate of mutations at these sites (Discussed in Chapter 1, section 1.3.5.1.3) thereby increasing the risk of development of tumours. This information could potentially impact the future of regulatory toxicology and environmental biomonitoring.

6.6. Gaps in knowledge and future work

1. The combination of MeDIP with zebrafish CGI tiling microarray in Chapter 3 and with HTS in Chapter 4 provided a valuable relative measurement of DNA methylation levels. An alternative approach would be to achieve an absolute level of DNA methylation especially for normal adult zebrafish liver.
2. In Chapter 4, it is possible that EDCs in association with other categories of compounds including genotoxic chemicals, such as PAHs are causing the changes observed in ST and HCA samples. However, in this study the existence of ample evidence strongly supported the role of EDCs in development of these tumours. The studies presented in this thesis provide a strong justification for expanding this work to chemical exposure studies in controlled laboratory environments. Furthermore, such studies could be used to investigate the possibility of transgenerational epigenetic inheritance. As mentioned in Chapter 1, gene expression is regulated through a highly interactive network of epigenetic and non-epigenetic factors. If environmental factors can cause a heritable change to the epigenome with phenotypic consequences in an individual organism, it is also probable that if these changes

occur in germ cells they could affect generations to come (Whitelaw and Whitelaw, 2008; Anway et al., 2005; Anway and Skinner, 2006). This is a phenomenon known as transgenerational epigenetic inheritance which is briefly described in Chapter 1 (section 1.3.5.2). Therefore, it is possible that the high prevalence of liver tumours, induced by environmental stressors in certain populations of flatfish dab, is epigenetically transmitted through multiple generations (Figure 6.1, modified from Turner, 2009).

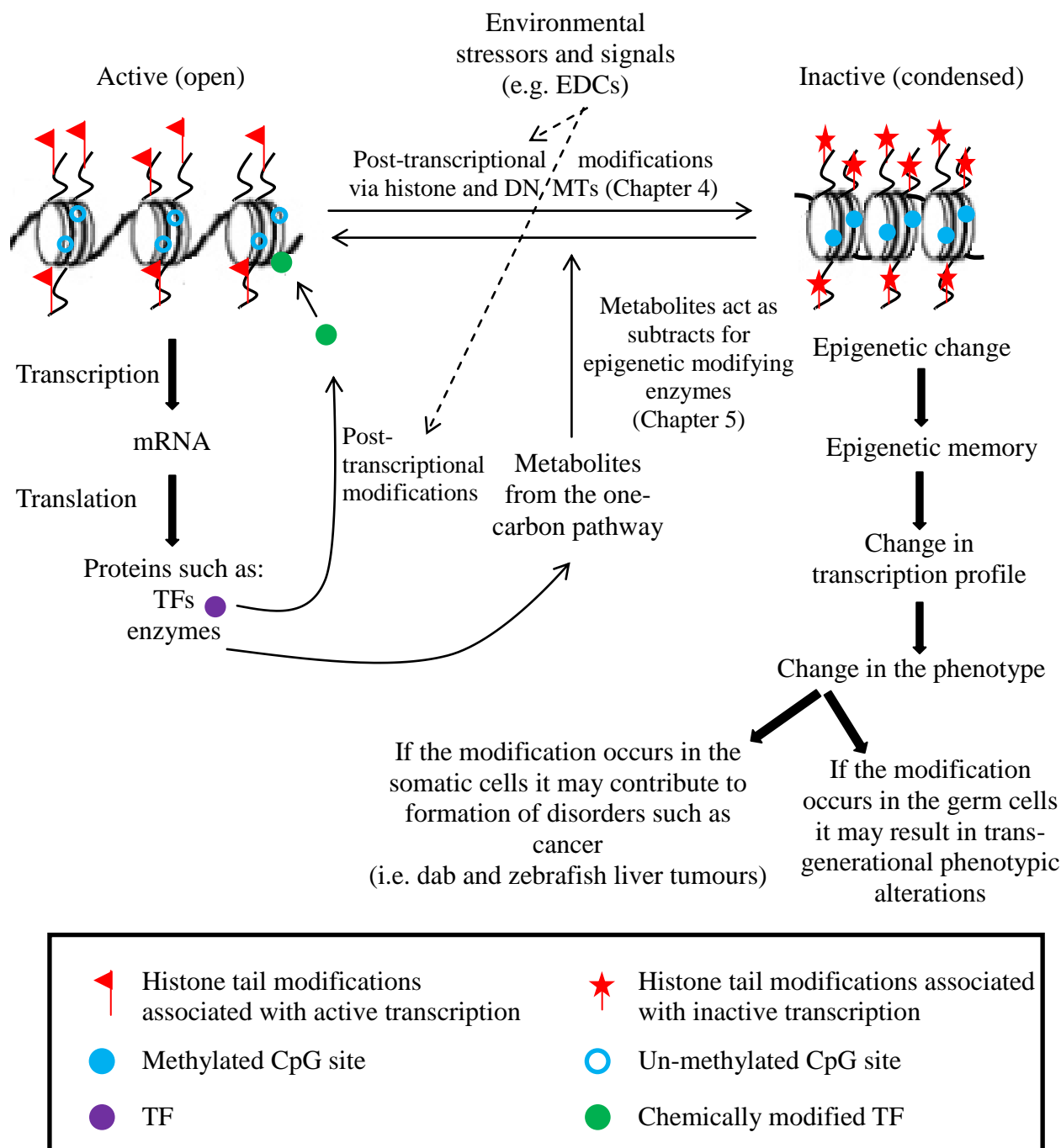


Figure 6.1. The complex network of epigenetic and non-epigenetic factors involved in the regulation of responses to environmental factors. The mechanism of response to environmental stimuli becomes more complex when considering that epigenetic modifications are not limited to modifications of histones and cytosine bases. Non-histone regulatory proteins such as transcription factors (TFs) undergo chemical modifications via histone modifying enzymes. This relationship becomes even more complex when considering that chemical modifications are interdependent and interactive. Furthermore, this figure demonstrates how a change in the epigenome induced by environmental stressors may be transmitted to multiple generations. This figure was modified from a review by Turner, 2009.

3. As mentioned in section 6.4, some of the changes observed in DNA methylation between various categories of samples could be due to differences in the ratios of various cell populations (e.g. hepatic cells, inflammatory cells, different stages of HCA tumours). A pure cell population can be achieved with techniques such as laser capture microdissection (LCMD). Furthermore, LCMD can be used to dissect cells at various stages of tumourigenesis. Such sampling could help to investigate the order of events (i.e. the occurrence of DNA hypomethylation at very early stages of tumourigenesis). The work presented in this thesis justifies the need for optimisation of methods for studying DNA methylation changes in LCMD samples.

4. In Chapter 5, the concentrations of the metabolites involved in the one-carbon cycle were quantified as peak area, and the relative change between the three categories of interest (HCA, ST, healthy dab liver) were investigated. This work can be followed up by using chemical standards to measure the absolute concentrations of these metabolites in dab tissue.

6.7. Concluding remarks

The work presented in this thesis demonstrated for the first time the involvement and alteration of DNA methylation in liver tumours of any aquatic species and provided a suitable strategy for studying DNA methylation in un-sequenced species. Furthermore, this work provided some evidence of a link between environmental stressors, epigenetic modifications and phenotypic end points. The data presented in this thesis demonstrated the need for consideration of the role of epigenetic mechanisms in regulation of responses of aquatic species to environmental toxicants which could potentially have a significant impact in procedures currently in place for environmental toxicity safety assessments and biomonitoring. Above all, it has opened a new research area in the field of aquatic biology and with it, brings many questions.

Chapter 7

References

Adrian, C., Creagh, E.M., Martin, S.J. (2001) Apoptosis-associated release of Smac/Diablo from mitochondria requires active caspases and is blocked by Bcl-2. *The EMBO Journal*, 20 (23): 6627-6638.

Akers, S.N., Odunsi, K., Karpf, A.R. (2010) Regulation of cancer germline antigen gene expression: Implications for cancer immunotherapy. *Future Oncology*, 6 (5): 717-732.

Amatruda, J.F., Shepard, J.L., Stern, H.M., Zon, L.I. (2002) Zebrafish as a cancer model system. *Cancer Cell*, 1 (3): 229-231.

Anderson, L., Hunter, C.L. (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Molecular and Cellular Proteomics*, 5 (4): 573-588.

Aniagu, S.O., Williams, T.D., Allen, Y., Katsiadaki, I., Chipman, J.K. (2008) Global genomic methylation levels in the liver and gonads of the three-spine stickleback (*Gasterosteus aculeatus*) after exposure to hexabromocyclododecane and 17-[beta] oestradiol. *Environment International*, 34 (3): 310-317.

Anway MD, Cupp AS, Uzumcu M, Skinner MK. (2005) Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science*, 308 (6): 1466–1469.

Anway MD, Memon MA, Uzumcu M, Skinner MK. (2006) Transgenerational effect of the endocrine disruptor vinclozolin on male spermatogenesis. *Journal of Andrology*, 2006, 27 (6):868–879.

Anway MD, Rekow SS, Skinner MK. (2008a) Transgenerational epigenetic programming of the embryonic testis transcriptome. *Genomics*, 91 (1): 30–40

Anway, M.D., Rekow, S.S., Skinner, M.K. (2008b) Comparative anti-androgenic actions of vinclozolin and flutamide on transgenerational adult onset disease and spermatogenesis. *Reproductive Toxicology*, 26 (2): 100-106.

Anway, M.D., Skinner, M.k. (2006) Epigenetic transgenerational actions of endocrine disruptors. *Endocrinology*, 147 (6): S43-S49.

Anway MD, Skinner MK. (2008) Transgenerational effects of the endocrine disruptor vinclozolin on the prostate transcriptome and adult onset disease. *The Prostate*, 68 (5): 517–529.

Archer, K.J., Mas, V.R., Maluf, D.G., Fisher, R.A. (2010) High-throughput assessment of CpG site methylation for distinguishing between HCV-cirrhosis and HCV-associated hepatocellular carcinoma. *Molecular Genetics and Genomics*, 283 (4): 341-349.

Arita, A., Costa, M. (2009) Epigenetics in metal carcinogenesis: Nickel, Arsenic, Chromium and Cadmium. *Metallomics*, 1 (3): 222-228.

- Armstrong, B., Doll, R. (1975) Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *International Journal of Cancer*, 15 (4): 617-631.
- Avner, P., Heard, E. (2001) X-Chromosome inactivation: Counting, choice and initiation. *Nature Review Genetics*, 2 (1): 59-67.
- Baccarelli, A., Bollati, V. (2009) Epigenetics and environmental chemicals. *Current Opinion in Paediatrics*, 21 (2): 243-251.
- Bachman, A.N., Phillips, J.M. and Goodman, J.I. (2006) Phenobarbital induces progressive patterns of GC-rich and gene-specific altered DNA methylation in the liver of tumour-prone B6C3F1 mice. *Toxicological Sciences*, 91 (2): 393-405.
- Bailey, G.S., Williams, D.E., Hendricks, J.D. (1996) Fish models for environmental carcinogenesis: The rainbow trout. *Environmental Health Perspectives*, 104 (1): 5-21.
- Barrett, J.C. (1993) Mechanisms of multistep carcinogenesis and carcinogen risk assessment. *Environmental Health Perspectives*, 100: 9-20.
- Barros, S.P., Offenbacher, S. (2009) Epigenetics: Connecting environment and genotype to phenotype and disease. *Journal of Dental Research*, 88 (5): 400-408.
- Bastow, R., Mylne, J.S., Lister, C., Lippman, Z., Martienssen, R.A., Dean, C. (2004) Vernalisation requires epigenetic silencing of FLC by histone methylation. *Nature*, 427 (6970): 164-167.
- Baylin, S.B., Esteller, M., Rountree, M.R., Bachman, K.E., Schuebel, K., Herman, J.G. (2001) Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Human Molecular Genetics*, 10 (7): 687-692.
- Bell, A.C., West, A.G., Felsenfeld, G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, 98 (3): 387-396.
- Benjamini, Y., Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57 (1): 289-300.
- Berghmans, S., Jette, C., Langenau, D., Hsu, K., Stewart, R., Look, T., Kanki, J.P. (2005) Making waves in cancer research: New models in the zebrafish. *Biotechniques*, 39 (2): 227-237.
- Bernstein, B.E., Meissner, A., Lander, E.S. (2007) The mammalian epigenome. *Cell*, 128 (4): 669-681.
- Bestor, T., Laudano, A., Mattaliano, R., Ingram, V. (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the

mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of Molecular Biology*, 203 (4): 971-983.

Bikkavilli, R.K., Malbon, C.C. (2009) Mitogen-activated protein kinases and Wnt/b-catenin signalling: Molecular conversations among signalling pathways. *Communicative and Integrative Biology*, 2 (1): 46-49.

Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes and Development*, 16 (1): 6-21.

Bird, A. (1995) Gene number, noise reduction and biological complexity. *Trends in Genetics*, 11 (3): 94-100.

Bird, A. (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research*, 8 (7): 1499-1504.

Bjornstrom, L., Sjoberg, M. (2005) Mechanisms of estrogen receptor signalling: Convergence of genomic and nongenomic actions on target genes. *Molecular Endocrinology*, 19 (4): 833-842.

Bollati, V., Baccarelli, A. (2010) Environmental epigenetics. *Heredity*, 105 (105): 112.

Bourchis, D., Xu, G.L., Lin, C.S., Bollman, B., Bestor, T.H. (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science*, 294 (5551): 2536-2539.

Bragiganda, V., Amiard-Triqueta, C., Parlierc, E., Bouryc, P., Marchandd, P., El Hourchb, M. (2006) Influence of biological and ecological factors on the bioaccumulation of polybrominated diphenyl ethers in aquatic food webs from French estuaries. *Science of the Total Environment*, 368 (2-3): 615-626.

Brosnan, J.T., Jacobs, R.L., Stead, L.M., Brosnan, M.E. (2004) Methylation demand: A key determinant of homocysteine metabolism. *Acta Biochimica Polonica*, 51 (2): 405-413.

Burgess-Beusse, B., Farrell, C., Gaszner, M., Litt, M., Mutskov, V., Recillas-Targa, F., Simpson, M., West, A., Felsenfeld, G. (2002) The insulation of genes from external enhancers and silencing chromatin. *Proceedings of the National Academy of Sciences*, 99 (4): 16433-16437.

Burn, J.E., Bagnall, D.J., Metzger, J.D., Dennis, E.S., Peacock, W.J. (1993) DNA methylation, vernalisation, and the initiation of flowering. *Proceedings of the National Academy of Science of the United States of America*, 90 (1): 287-291.

Butler, J.S., Lee, J.H., Skalnik, D.G. (2008) CFP1 interacts with DNMT1 independently of association with the Setd1 histone H3K4 methyltransferase complexes. *DNA and Cell Biology*, 27 (10): 533-543.

Caudill, M.A., Wang, J.C., Melnyk, S., Pogribny, I.P., Jernigan, S., Collins, M.D., Santos-Guzman, J., Swendseid, M.E., Cogger, E.A., James, S.J. (2001) Intracellular S-

adenosylhomocysteine concentrations predict global DNA hypomethylation in tissues of methyl-deficient cystathionine beta-synthase heterozygous Mice. *The Journal of Nutrition*, 131 (11): 2811-2818.

Cha, C., DeMatteo, R.P. (2005) Molecular mechanisms in hepatocellular carcinoma development. Best Practice and Research. *Clinical Gastroenterology*, 19 (1): 25-37.

Chan, T.A., Glockner, S., Yi, J.M., Chen, W., Van Neste, L., Cope, L., Herman, J.G., Velculescu, V., Schuebel, K.E., Ahuja, N., Baylin, S.B. (2008) Convergence of mutation and epigenetic alterations identifies common genes in cancer that predict for poor prognosis. *PLoS Medicine*, 5 (5): 823-838.

Chen, N.C., Yang, F., Capecci, L.M., Gu, Z., Schafer, A.I., Durante, W., Yang, X., Wang, H. (2010) Regulation of homocysteine metabolism and methylation in human and mouse tissues. *The Journal of the Federation of American Societies for Experimental Biology*, 24 (8): 2804-2817.

Cheng, X., Blumenthal, R.M. (2008) Mammalian DNA methyltransferases: A structural perspective. *Structure*, 16 (3): 341-350.

Choi, S.W., Mason, J.B. (2002) Folate status: Effects on pathways of colorectal carcinogenesis. *Journal of Nutrition*, 132 (8): 2413S-2418S.

Christensen, B.C., Andres Houseman, E., Marsit, C.J., Zheng, S., Wrensch, M.R., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Padbury, J.F., Bueno, R., Sugarbaker, D.J., Yeh, R., Wiencke, J.K., Kelsey, K.T. (2009) Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genetics*, 5 (8): 1-13.

Clark, I.M., Swingler, T.E., Sampieri, C.L., Edwards, D.R. (2008) The regulation of matrix metalloproteinases and their inhibitors. *The International Journal of Biochemistry and Cell Biology*, 40 (6-7): 1326-1378.

Cohen, R., Chalifa-Caspi, V., Williams, T.D., Auslander, M., George, S.G., Chipman, J.K., Tom, M. (2007) Estimating the efficiency of fish cross-species cDNA microarray hybridisation. *Marine Biotechnology*, 9 (4): 491-499.

Coleman, M.L., Sahai, E.A., Yeo, M., Bosch, M., Dewart, A., Olson, M.F. (2001) Membrane blebbing during apoptosis results from caspase-mediated activation of ROCK 1. *Nature Cell Biology*, 3 (4): 339-346.

Conesa, A., Gotz, S., Garcia-Gomez, J., Terol, J., Talon, M., Robles, M. (2005) Blat2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21 (18): 3674-3676.

Contractor, R.G., Foran, C.M., Li, S., Willett, K.L. (2004) Evidence of gender- and tissue-specific promoter methylation and the potential for ethinylestradiol induced changes in Japanese medaka (*Oryzias latipes*) estrogen receptor and aromatase genes. *Environmental Health*, 67 (1): 1-22.

Coulondre, C., Miller, J.H., Farabaugh, P.J., Gilbert, W. (1978) Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature*, 274 (5673): 775-780.

Counts, J.L., Goodman, J.I. (1995) Hypomethylation of DNA: A nongenotoxic mechanism involved in tumour promotion. *Toxicology Letters*, 82-83: 663-672.

Cox, D.M., Zhong, F., Du, M., Duchoslav, E., Sakuma, T., McDermott, J.C. (2005) Multiple reaction monitoring as a method for identifying protein posttranslational modifications. *Journal of Biomolecular Techniques*, 16 (2): 83-90.

Cravo, M., Pinto, R., Fidalgo, P., Chaves, P., Gloria, L., Nobre-Leitao, C., Costa Mira, F. (1996) Global DNA hypomethylation occurs in the early stages of intestinal type gastric carcinoma. *Gut*, 39 (3): 434-438.

Csankovszki, G., Nagy, A., Jaenisch, R. (2001) Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. *The Journal of Cell Biology*, 153 (4): 773-783.

Cui, H., Cruz-Correa, M., Giardiello, F.M., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., He, X., Powe, N.R., Feinberg, A.P. (2003) Loss of IGF2 imprinting: A potential marker of colorectal cancer risk. *Science*, 299 (5613): 1753-1755.

Cui, P., Zhang, L., Lin, Q., Ding, F., Xin, C., Fang, X., Hu, S., Yu, J. (2010) A novel mechanism of epigenetic regulation: Nucleosome-space occupancy. *Biochemical and Biophysical Research Communications*, 391 (1): 884-889.

Davis, S.R., Quinlivan, E.P., Shelnutt, K.P., Maneval, D.R., Ghandour, H., Capdevila, A., Coats, B.S., Wagner, C., Selhub, J., Bailey, L.B., Shuster, J.J., Stacpoole, P.W., Gregory, J.F. (2005) The methylenetetrahydrofolate reductase 677C→T polymorphism and dietary folate restriction affect plasma one-carbon metabolites and red blood cell folate concentrations and distribution in women. *The journal of Nutrition*, 135 (5): 1040-1044.

De Bont, R., Van Larebeke, N. (2004) Endogenous DNA damage in humans: A review of quantitative data. *Mutagenesis*, 19 (3): 169-185.

De Camargo, J.L.V., Punyarit, P., Newberne, P.M. (1985) Early stages of nodular transformation of the B6C3F1 mouse liver induced by choline deficiency. *Toxicologic Pathology*, 13 (1): 10-17.

De Smet, C., Lorient, A. (2010) DNA hypomethylation in cancer. Epigenetic scars of a neoplastic journey. *Epigenetics*, 5 (3): 206-213.

De Smet, C., Lorient, A., Boon, T. (2004) Promoter-dependent mechanism leading to selective hypomethylation within the 5' region of gene MAGE-A1 in tumour cells. *Molecular and Cellular Biology*, 24 (11): 4781-4790.

DeLong, C.J., Hicks, A.M., Cui, Z. (2002) Disruption of choline methyl group donation for phosphatidylethanolamine methylation in hepatocarcinoma cells. *The Journal of Biological Chemistry*, 277 (19): 17217-17225.

Dhillon, A.S., Hagan, S., Rath, O., Kolch, W (2007) MAP kinase signalling pathways in cancer. *Oncogene*, 26 (22): 3279-3290.

Diab, A.M., Williams, T.D., Sabine, V.S., Chipman, J.K., George, S. G. (2008) The GENIPOL European flounder *Platichthys flesus* L. toxicogenomics microarray: Application for investigation of the response to furunculosis vaccination. *Journal of Fish Biology*, 72 (9): 2154-2169.

Doerfler, W., Kruczek, I., Eick, D., Vardimon, L., Kron, B. (1983) DNA methylation and gene activity: The adenovirus system as a model. *Cold Spring Harbor Symposia on Quantitative Biology*, 47: 593-603.

Dolcet, X., Liobet, D., Pallares, J., Matias-Guiu, X. (2005) NF- κ B in development and progression of human cancer. *Virchows Archiv: An International Journal of Pathology*, 446 (5): 475-482.

Dolinoy, D.C. (2008) The agouti mouse model: An epigenetic biosensor for nutritional and environmental alterations on the foetal epigenome. *Nutrition reviews*, 66 (suppl 1): 1-8.

Dolinoy, D.C., Das, R., Weidman, J.R., Jirtle, R.L. (2007) Metastable epialleles, imprinting, and the foetal origins of adult diseases. *Paediatric Research*, 61 (5): 30-37.

Dolinoy, D.C., Wiedman, J., Waterland, R., Jirtle, R.L. (2006) Maternal genistein alters coat colour and protects Avy mouse offspring from obesity by modifying the foetal epigenome. *Environmental Health Perspectives*, 114 (4): 567-572.

Doll, R., Peto, R. (1981) The causes of cancer: Quantitative estimates of avoidable risks of cancer in the United States today. *Journal of the National Cancer Institute*, 66 (6): 1191-1308.

Down, T.A., Rakyan, V.K., Turner, D.J., Flicek, P., Li, H., Kulesha, E., Graf, S., Johnson, N., Herrero, J., Tomazou, E.M., Thorne, N.P., Backdahl, L., Herberth, M., Howe, K.L., Jackson, D.K., Miretti, M.M., Marioni, J.C., Birney, E., Hubbard, T.J.P., Durbin, R., Tavare, S., Beck, S. (2008) A bayesian deconvolution strategy for immunoprecipitation based DNA methylome analysis. *Nature Biotechnology*, 26 (7): 779-785.

Durand, P., Prost, M., Loreau, N., Lussier-Cacan, S., Blache, D. (2001) Impaired homocysteine metabolism and atherothrombotic disease. *Laboratory Investigation*, 81 (5): 645-672.

Duthie, S. (2011) Folate and cancer: How DNA damage, repair and methylation impact on colon carcinogenesis. *Journal of Inherited Metabolic Disease*, 34 (1): 101-109.

- Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V.K., Attwood, J., Burger, M., Burton, J., Cox, T.V., Davies, R., Down, T.A., Haefliger, C., Horton, R., Howe, K., Jackson, D.K., Kunde, J., Koenig, C., Liddle, J., Niblett, D., Otto, T., Pettett, R., Seemann, S., Thompson, C., West, T., Rogers, J., Olek, A., Berlin, K., Beck, S. (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature Genetics*, 38 (12): 1378-1385.
- Egeblad, M., Werb, Z. (2002) New functions for the matrix metalloproteinases in cancer progression. *Nature reviews Cancer*, 2 (3): 161-174.
- Egger, G., Liang, G., Appricio, A., Jones, P.A. (2004) Epigenetics in human disease and prospects for epigenetic therapy. *Nature*, 429 (6990): 457-460.
- Ehrlich, M. (2009) DNA hypomethylation in cancer cells. *Epigenomics*, 1 (2): 239-259.
- Ehrlich, M. (2005) The controversial denouement of vertebrate DNA methylation research. *Biochemistry (Moscow) Journal*, 70 (5): 690-698.
- Emmert, J.L., Garrow, T.A., Baker, D.H. (1996) Hepatic betaine-homocysteine methyltransferase activity in the chicken is influenced by dietary intake of sulfur amino acids, choline and betaine. *The Journal of Nutrition*, 126 (8): 2050-2058.
- Esteller, M. (2008) Epigenetics in cancer. *The New England Journal of Medicine*, 358 (11): 1148-1159.
- Esteller, M. (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nature Review Genetics*, 8 (4): 286-298.
- Esteller, M., Corn, P.G., Baylin, S.B., Herman, J.G. (2001) A gene hypermethylation profile of human cancer. *Cancer Research*, 61 (8): 3225-3229.
- Esteller, M., Herman, J.G. (2002) Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *Journal of Pathology*, 196 (1): 1-7.
- Fang, Y., Huang, H., Juan, H. (2008) MeInfoText: Associated gene methylation and cancer information from text mining. *BMC Bioinformatics*, 9 (22): 1-9.
- Feinberg, A.P. (2001) Cancer epigenetics takes center stage. *Proceedings of the National Academy of Science of the United States of America*, 98 (2): 392-394.
- Feinberg, A.P. (2004) The epigenetics of cancer etiology. *Seminars in Cancer Biology*, 14 (6): 427-432.
- Feinberg, A.P., Ohlsson, R., Henikoff, S. (2006) The epigenetic progenitor origin of human cancer. *Nature Review Genetics*, 7 (1): 21-33.
- Feist, S.W., Lang, T., Stentiford, G.D., Koehler, A. (2004) Biological effects of contaminants: use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder

(*Platichthys flesus* L.) for monitoring. *ICES Techniques in Marine Environmental Sciences*, 38 (8-9): 1-42.

Feitsma, H., Cuppen, E. (2008) Zebrafish as a cancer model. *Molecular Cancer Research*, 6 (5): 685-694.

Feldser, D., Agani, F., Iyer, N.V., Pak, B., Ferreira, G., Semenza, G.L (1999) Reciprocal positive regulation of hypoxia-inducible factor 1 α and insulin-like growth factor 2. *Cancer Research*, 59 (16): 3915-3918.

Fenga, S., Cokus, S.J., Zhang, X., Chen, P., Bostick, M., Golle, M.G., Hetzel, J., Jaine, J., Strauss, S.H., Halpern, M.E., Ukomadug, C., Sadler, K.C., Pradhani, S., Pellegrini, M., Jacobsen, S.E. (2010) Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Science of the United States of America*, 107 (19): 8689-8694.

Firestein, R., Bass, A.J., Young Kim, S., Dunn, I.F., Silver, S.J., Guney, I., Freed, E., Ligon, A.H., Vena, N., Ogino, S., Chheda, M.G., Tamayo, P., Finn, S., Shrestha, Y., Boehm, J.S., Jain, S., Bojarski, E., Mermel, C., Barretina, J., Chan, J.A., Baselga, J., Tabernero, J., Root, D.E., Fuchs, C.S., Loda, M., Shivdasani, R.A., Meyerson, M., Hahn, W.C. (2008) CDK8 is a colorectal cancer oncogene that regulates β -catenin activity. *Nature*, 455 (7212): 547-551.

Fox, I.H., Kelley, W.N. (1978) The role of adenosine and 2'-deoxyadenosine in mammalian cells. *Annual Review of Biochemistry*, 47: 655-686.

Fraga, M.F., Ballester, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudose, J.C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T.D., Wu, Y-Z., Plass, C., Esteller, M. (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Science of the United States of America*, 102 (30): 10604-10609.

Franklin, T.M., Lee, J., Kohler, A., Chipman, J.K. (2000) Analysis of mutations in the p53 tumour suppressor gene and Ki- and Ha-ras proto-oncogenes in hepatic tumours of European flounder (*Platichthys flesus*). *Marine Environmental Research*, 50 (1-5): 251-255.

Gama-Sosa, M.A., Slagel, V.A., Trewyn, R.W., Oxenhandler, R., Kuo, K.C., Gehrke, C.W., Ehrlich, M. (1983) The 5-methylcytosine content of DNA from human tumours. *Nucleic Acids Research*, 11 (19): 6883-6894.

Gao, W., Kondo, Y., Shen, L., Shimizu, Y., Sano, T., Yamao, K., Natsume, A., Goto, Y., Ito, M., Murakami, H., Osada, H., Zhang, J., Issa, J.J., Sekido, Y. (2008) Variable DNA methylation patterns associated with progression of disease in hepatocellular carcinomas. *Carcinogenesis*, 29 (10): 1901-1910.

Gardiner-Garden, M., Frommer, M. (1987) CpG islands in vertebrate genome. *Journal of Molecular Biology*, 196 (2): 261-282.

Gaudet, F., Hodgson, J.G, Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., Jaenisch, R. (2003) Induction of tumours in mice by genomic hypomethylation. *Science*, 300 (5618): 489-492.

Ghoshal, A.K., Farber, E. (1984) The induction of liver cancer by dietary deficiency of choline and methionine without added carcinogens. *Carcinogenesis*, 5 (10): 1367-1370.

Giovannucci, E., Stampfer, M.J., Colditz, G.A., Rimm, E.B., Trichopoulos, D., Rosner, B.A., Speizer, F.E., Willett, W.C. (1993) Folate, methionine, and alcohol intake and risk of colorectal adenoma. *Journal of The National Cancer Institute*, 85 (11): 875-883.

Girault, I., Tozlu, S., Lidereau, R., Bieche, I. (2003) Expression analysis of DNA methyltransferases 1, 3A, and 3B in sporadic breast carcinomas. *Clinical Cancer Research*, 9 (12): 4415-4422.

Godmann, M., Lambrot, R., Kimmins, S. (2009) The dynamic epigenetic program in male germ cells: Its role in spermatogenesis, testis cancer, and its response to the environment. *Microscopy Research and Technique*, 72 (8): 603-619.

Goodman, J.I., Augustine, K.A., Cunningham, M.L., Dixon, D., Dragan, Y.P., Falls, J.G., Rasoulpour, R.J., Sills, R.C., Storer, R.D., Wolf, D.C., Pettit, S.D. (2010) What do we need to know prior to thinking about incorporating an epigenetic evaluation into safety assessments?. *Toxicological Sciences*, 116 (2): 375-381.

Goll, M.G., Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. *Annual Review of Biochemistry*, 74: 481-514.

Gotz, S., Garcia-Gomez, J., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J., Conesa, A. (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, 36 (10): 3420-3435.

Grabher, C., Look, A.T. (2006) Fishing for cancer models. *Nature Biotechnology*, 24 (1): 45-46.

Griffin, J.L., Shockcor, J.P. (2004) Metabolic profiles of cancer cells. *Nature Reviews Cancer*, 4 (7): 551-561.

Gronbaek, K., Hother, C., Jones, P.A. (2007) Epigenetic changes in cancer. *Acta Pathologica Microbiologica et Immunologica Scandinavica*, 115 (10): 1039-1059.

Guerrero-Bosagna, C., Sabat, P., Valladares, L. (2005) Environmental signalling and evolutionary change: Can exposure of pregnant mammals to environmental estrogens lead to epigenetically induced evolutionary changes in embryos? *Evolution and Development*, 7 (4): 341-350.

Hamaguchi, T., Norio, I., Ryouichi, T., Yoshihiko, H., Takanobu, M., Michihisa, I., Yoshihiro, T., Kazuhiko, S., Motonari, T., Takao, T., Masaaki, O. (2008) Glycolysis module

activated by hypoxia-inducible factor 1 α is related to the aggressive phenotype of hepatocellular carcinoma. *International Journal of Oncology*, 33 (4): 725-731.

Hanahan, D., Weinberg, R.A. (2000) The hallmarks of cancer. *Cell*, 100 (1): 57-70.

Hanahan, D., Weinberg, R.A. (2011) Hallmarks of cancer: The next generation. *Cell*, 144 (5): 646-674.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R., S., Levorse, J.M., Tilghman, S.M. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, 405 (6785): 486-489.

Hawkins, W.E., Walker, W., Overstreet, R.M., Lytle, T.F., Lytle, J.S. (1988) Dose-related carcinogenic effects of water-borne benzo[α]pyrene on livers of two small fish species. *Ecotoxicology and Environmental Safety*, 16 (3): 219-231.

He, Y., Michaels, S.D., Amasino, R.M. (2003) Regulation of flowering time by histone acetylation in Arabidopsis. *Science*, 302 (5651): 1751-1754.

Heinzmann, S.S., Brown, I.J., Chan, Q., Bictash, M., Dumas, M., Kochhar, S., Stamler, J., Holmes, E., Elliott, P., Nicholson, J.K. (2010) Metabolic profiling strategy for discovery of nutritional biomarkers: Proline betaine as a marker of citrus consumption. *The American Journal of Clinical Nutrition*, 92 (2): 436-443.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V.A., Bird, A. (2001) Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes and Development*, 15 (6): 710-723.

Herceg, Z., Vaissiere, T. (2011) Epigenetic mechanisms and cancer. An interface between the environment and the genome. *Epigenetics*, 6 (7): 804-819.

Herrera, L.A., Prada, D., Andonegui, M.A., Dueñas-González, A. (2008) The Epigenetic Origin of Aneuploidy. *Current Genomics*, 9 (1): 43-50.

Ho, S., Tang, W., de Frausto, J.B., Prins, G.S. (2006) Developmental exposure to estradiol and bisphenol A increases susceptibility to prostate carcinogenesis and epigenetically regulates phosphodiesterase type 4 variant 4. *Cancer Research*, 66 (11): 5624-5632.

Hoeflich, A., Reisinger, R., Lahm, H., Kiess, W., Blum, W.F., Kolb, H.J., Weber, M.M., Wolf, E. (2001) Insulin-like growth factor-binding protein 2 in tumourigenesis: Protector or promoter? *Cancer Research*, 61 (24): 8601-8610.

Hore, T.A., Rapkins, R.W., Marshall Graves, J.A. (2007) Construction and evolution of imprinted loci in mammals. *Trends in Genetics*, 23 (9): 440-448.

Hu, W., Johnson, H., Shu, H. (2000) Activation of NF- κ B by FADD, casper, and caspase-8. *Journal of Biological Chemistry*, 275 (15): 10838-10844.

- Huang, Y., Pastor, W.A., Shen, Y., Tahiliani, M., Liu, D.R., Rao, A. (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One*, 5 (1): 1-9.
- Illingworth, R.S., Bird, A.P. (2009) CpG islands-‘A rough guide’. *Federation of European Biochemical societies Letters*, 583 (11): 1713-1720.
- Inawaka K, Kawabe M, Takahashi S, Doi Y, Tomigahara Y, Tarui H, Abe J, Kawamura S, Shirai T. (2009) Maternal exposure to anti-androgenic compounds, vinclozolin, flutamide and procymidone, has no effects on spermatogenesis and DNA methylation in male rats of subsequent generations. *Toxicology and Applied Pharmacology*, 237 (2): 178–187.
- Inman, B.A., Harel, F., Audet, J.F., Meyer, F., Douville, P., Fradet, Y., Lacombe, L. (2005) Insulin-like growth factor binding protein 2: An androgen-dependent predictor of prostate cancer survival. *European Urology*, 47 (5): 695-702.
- Jabbari, K., Bernardi, G. (2004) Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene*, 333: 143-149.
- Jaenisch, R., Bird, A. (2003) Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nature Genetics*, 33 (suppl): 245-254.
- James, S.J., Melnyk, S., Pogribna, M., Pogribny, I.P., Caudill, M.A. (2002) Elevation in S-adenosylhomocysteine and DNA hypomethylation: Potential epigenetic mechanism for homocysteine-related pathology. *The Journal of Nutrition*, 132 (8 Suppl): 2361S-2366S.
- Jiang, Z., Liang, Q., Luo, G., Huc, P., Li, P., Wang, Y (2009) HPLC–electrospray tandem mass spectrometry for simultaneous quantitation of eight plasma amino thiols: Application to studies of diabetic nephropathy. *Talanta*, 77 (4): 1279-1284.
- Jirtle, R.L., Skinner, M.K. (2007) Environmental epigenomics and disease susceptibility. *Nature Review Genetics*, 8 (4): 253-262.
- Jones, P.A. (1999) The DNA methylation paradox. *Trends in Genetics*, 15 (1): 34-37.
- Jones, P.A., Liang, G. (2009) Rethinking how DNA methylation patterns are maintained. *Nature Review Genetics*, 10 (11): 805-811.
- Jones, P.A., Takai, D. (2001) The role of DNA methylation in mammalian epigenetics. *Science*, 293 (5532): 1068-1070.
- Kangaspeska, S., Stride, B., Metivier, R., Polycarpou-Schwarz, M., Ibberson, D., Carmouche, R.P., Benes, V., Gannon, F., Reid, G. (2008) Transient cyclical methylation of promoter DNA. *Nature*, 452 (7183): 112-116.
- Karp, G. C. (2000) Cell and molecular biology: Concepts and experiments. Fourth ed. *John Wiley and Sons Ltd*, 1-864.

Kelly, C., Smallbone, K., Brady, M. (2008) Tumour glycolysis: The many faces of HIF. *Journal of Theoretical Biology*, 254 (2): 508-513.

Khudoley, W. (1984) Use of aquarium fish, *Danio rerio* and *Poecilia reticulata*, as test species for evaluation of nitrosamine carcinogenicity. *National Cancer Institute Monograph*, 65: 65-70.

Klaunig, J.E., Kamendulis, L.M., Xu, Y. (2000) Epigenetic mechanisms of chemical carcinogenesis. *Human and Experimental Toxicology*, 19 (10): 543-555.

Klug, M., Heinz, S., Gebhard, C., Schwarzfischer, L., Krause, S.W., Andreesen, R., Rehli, M. (2010) Active DNA demethylation in human postmitotic cells correlates with activating histone modifications, but not transcription levels. *Genome Biology*, 11 (6): 1-11.

Kohli, M., Yu, J., Seaman, C., Bardelli, A., Kinzler, K.W., Vogelstein, B., Lengauer, C., Zhang, L. (2004) SMAC/Diablo-dependent apoptosis induced by non-steroidal anti-inflammation drugs (NSAIDs) in colon cancer cells. *Proceedings of the National Academy of Science of the United States of America*, 101 (48): 16897-16902.

Kohlmeier, M., da Costa, K., Fischer, L.M., Zeisel, S.H. (2005) Genetic variation of folate-mediated one-carbon transfer pathway predicts susceptibility to choline deficiency in humans. *Proceedings of the National Academy of Science of the United States of America*, 102 (44): 16025-16030.

Korfmacher, W.A. (2005) Foundation review: Principles and applications of LC-MS in new drug discovery. *Drug Discovery Today*, 10 (20): 1357-1367.

Krijt, J., Duta, A., Kozich, V. (2009) Determination of S-adenosylmethionine and S-adenosylhomocysteine by LC-MS/MS and evaluation of their stability in mice tissues. *Journal of Chromatography B*, 877 (22): 2061-2066.

Kumar, R., Srivastava, R., Singh, R.K., Surolia, A., Rao, D.N. (2008) Activation and inhibition of DNA methyltransferases by S-adenosyl-L-homocysteine analogues. *Bioorganic and Medicinal Chemistry*, 16 (5): 2276-2285.

Laird, P.W., Jaenisch, R. (1996) The role of DNA methylation in cancer genetics and epigenetics. *Annual Review of Genetics*, 30: 441-464.

Laird, P.W., Jaenisch, R. (1994) DNA methylation and cancer. *Human Molecular Genetics*, 3: 1487-1495.

Lam, H.L., Wu, Y.L., Vega, V.B., Miller, L.D., Spitsbergen, J., Tong, Y., Zhan, H., Govindarajan, K.R., Lee, S., Mathavan, S., Krishna Murthy, K.R., Buhler, D.R., Liu, E.T., Gong, Z. (2006) Conservation of gene expression signatures between zebrafish and human liver tumours and tumour progression. *Nature Biotechnology*, 24 (1): 73-75.

Lam, S.H., Gong, Z. (2006) Modelling liver cancer using zebrafish. A comparative oncogenomics approach. *Cell Cycle*, 5 (6): 573-577.

- Lawrence, C., Sanders, G.E., Varga, Z.M., Baumann, D.P., Freeman, A., Baur, B., Francis, M. (2009) Regulatory compliance and the zebrafish. *Zebrafish*, 6 (4): 453-456.
- Lee, T.F., Zhai, J., Meyers, B.C. (2010) Conservation and divergence in eukaryotic DNA methylation. *Proceedings of the National Academy of Science of the United States of America*, 107 (20): 9027-9028.
- Legler, J. (2010) Epigenetics: An emerging field in environmental toxicology. *Integrated Environmental Assessment and Management*, 6 (2): 314-315.
- Lempiainen, H., Muller, A., Brasa, S., Teo, S., Roloff, T.C., Morawiec, L., Zamurovic, N., Vicart, A., Funhoff, E., Couttet, P., Schubeler, D., Grenet, O., Marlowe, J., Moggs, J., Terranova, R. (2011) Phenobarbital mediates an epigenetic switch at the constitutive androstane receptor (CAR) target gene Cyp2b10 in the liver of B6C3F1 mice. *PLoS One*, 6 (3): 1-14.
- Lever, M., Slow, S. (2010) The clinical significance of betaine, an osmolyte with a key role in methyl group metabolism. *Clinical Biochemistry*, 43 (9):732-744.
- Lewin, J., Schmitt, A.O., Adorjan, P., Hildmann, T., Piepenbrock, C. (2004) Quantitative DNA methylation analysis based on four-dye trace data from direct sequencing of PCR amplicates. *Bioinformatics*, 20 (17): 1-8.
- Li, E. (2002) Chromatin modification and epigenetic reprogramming in mammalian development. *Nature Review Genetics*, 3 (9): 662-673.
- Li, L.C., Dahiya, R. (2002) MethPrimer: Designing primers for methylation PCRs. *Bioinformatics*, 18 (11): 1427-1431.
- Li, N., Ye, M., Li, Y., Yan, Z., Butcher, L.M., Sun, J., Han, X., Chen, Q., Zhang, X., Wang, J. (2010) Whole genome DNA methylation analysis based on high throughput sequencing technology. *Methods*, 52 (3): 203-213.
- Li, S., Hursting, S.D., Davis, B.J., McLachlan, J.A., Barrett, J.C. (2003) Environmental exposure, DNA methylation, and gene regulation: lessons from diethylstilbesterol-induced cancers. *Annals of the New York Academy of Sciences*, 983: 161-169.
- Liang, G., Salem, C.E., Yu, M.C., Nguyen, H.D., Gonzales, F.E., Nguyen, T.T., Nichols, P.w., Jones, P.A. (1998) DNA methylation differences associated with tumour tissues identified by genome scanning analysis. *Genomics*, 53 (3): 260-268.
- Liang, X., Liang, Q., Xia, J., Wang, Y., Hu, P., Wang, Y., Zheng, X., Zhang, T., Luo, G. (2009) Simultaneous determination of sixteen metabolites related to neural tube defects in maternal serum by liquid chromatography coupling with electrospray tandem mass spectrometry. *Talanta*, 78 (4-5): 1246-1252.
- Lieber, C.S., Packer, L. (2002) S-adenosylmethionine: molecular, biological, and clinical aspects--an introduction. *The American Journal of Clinical Nutrition*, 76 (5): 1148-1150.

- Lindesjoo, E., Thulin, J. (1994) Histopathology of skin and gills of fish in pulp mill effluents. *Disease of aquatic organisms*, 18 (2): 81-93.
- Livak, K.J., Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods*, 25 (4): 402-408.
- Lo, P., Sukumar, S. (2008) Epigenomics and breast cancer. *Pharmacogenomics*, 9 (12): 1879-1902.
- Locker, J., Reddy, T.V., Lombardi, B. (1986) DNA methylation and hepatocarcinogenesis in rats fed a choline-devoid diet. *Carcinogenesis*, 7 (8): 1309-1312.
- Loeb, L.A., Harris, C.C. (2008) Advances in chemical carcinogenesis: A historical review and prospective. *Cancer Research*, 68 (17): 6863-6872.
- Lopez, J., Percharde, M., Coley, H.M., Webb, A., Crook, T. (2009) The context and potential of epigenetics in oncology. *British Journal of Cancer*, 100 (4): 571-577.
- Loriot, A., De Plaen, E., Boon, T., De Smet, C. (2006) Transient down-regulation of DNMT1 methyltransferase leads to activation and stable hypomethylation of MAGE-A1 in melanoma cells. *The Journal of Biological Chemistry*, 281 (15): 10118-10126.
- Lu, S.C. (2000) S-adenosylmethionine. *The International Journal of Biochemistry and Cell Biology*, 32 (4): 391-395.
- Lujan, R.J., Capote, F.P., Marinas, A., de Castro, M.D.L (2008) Liquid chromatography/triple quadrupole tandem mass spectrometry with multiple reaction monitoring for optimal selection of transitions to evaluate nutraceuticals from olive-tree materials. *Rapid Communications in Mass Spectrometry: RCM*, 22 (6): 855-864.
- Lyons, B.P., Stentiford, G.D., Bignell, J., Goodsir, F., Sivyier, D.B., Devlin, M.J., Lowe, D., Beesley, A., Pascoe, C.K., Moore, M.N., Garnacho, E. (2006) A biological effects monitoring survey of Cardigan Bay using flatfish histopathology, cellular biomarkers and sediment bioassays: findings of the Prince Madog prize 2003. *Marine Environmental Research*, 62: s342-s346.
- MacKay, A.B., Mhanni, A.A., McGowan, R.A., Krone, P.H. (2007) Immunological detection of changes in genomic DNA methylation during early zebrafish development. *Genome*, 50 (8): 778-785.
- Macleod, D., Clark, V.H., Bird, A. (1999) Absence of genome-wide changes in DNA methylation during development of the zebrafish. *Nature Genetics*, 23 (2): 139-140.
- Mally, A., Chipman, J.K. (2002) Non-genotoxic carcinogens: Early effects on gap junctions, cell proliferation and apoptosis in the rat. *Toxicology*, 180 (3): 233-248.
- Mamoune, A., Luo, J., Lauffenburger, D.A., Wells, A. (2003) Calpain-2 as a target for limiting prostate cancer invasion. *Cancer Research*, 63 (15): 4632-4640.

- Mann, B., Gelos, M., Siedow, A., Hanski, M.L., Gratchev, A., Ilyas, M., Bodmer, W.f., Moyer, M.P., Riecken, E.O., Buhr, H.J., Hanski, C. (1999) Target genes of β -catenin-T cell-factor/lymphoid-enhancer-factor signalling in human colorectal carcinomas. *Proceedings of the National Academy of Science of the United States of America*, 96 (4): 1603-1608.
- Martin, C.C., Laforest, L., Akimenko, M., Ekker, M. (1999) A role for DNA methylation in gastrulation and somite patterning. *Developmental Biology*, 206 (2): 189-205.
- Martinez-Zamudio, R., Ha, H.C. (2011) Environmental epigenetics in metal exposure. *Epigenetics*, 6 (7): 820-827.
- Masahito, P., Ishikawa, T., Sugano, H. (1988) Fish tumours and their importance in cancer research. *Japanese Journal of Cancer Research*, 79 (5): 545-555.
- Massicotte, R., Whitelaw, E., Angers, B. (2011) DNA methylation. A source of random variation in natural populations. *Epigenetics*, 6 (4): 421-427.
- McLafferty, F.W. (1981) Tandem Mass Spectrometry. *Science*, 214 (4518): 280-287.
- Melnyk, S., Pogribna, M., Pogribny, I.P., Yi, P. and James, S.J. (2000) Measurement of plasma and intracellular *S*-adenosylmethionine and *S*-adenosylhomocysteine utilizing coulometric electrochemical detection: Alterations with plasma homocysteine and pyridoxal 5'-phosphate concentrations.. *Clinical Chemistry*, 46 (2): 265-272.
- Mhanni, A.A., McGowan, R.A. (2004) Global changes in genomic methylation levels during early development of the zebrafish embryo. *Development Genes and Evolution*, 214 (8): 412-417.
- Michel, V., Yuan, Z., RamsuBir, S., Bakovic, M. (2006) Choline transport for phospholipid synthesis. *Experimental Biology and Medicine*, 231 (5): 490-504.
- Mill, J., Yazdanpanah, S., Guckel, E., Ziegler, S., Kaminsky, Z., Petronis, A. (2006) Whole genome amplification of sodium bisulfite-treated DNA allows the accurate estimate of methylated cytosine density in limited DNA recourses. *Biotechniques*, 41 (5): 603-607.
- Miller, J. (1970) Carcinogenesis by chemicals: An overview- G. H. A. Clowes memorial lecture. *Cancer Research*, 30 (3): 559-576.
- Mirbahai, L., Williams, T.D., Zhan, H., Gong, Z., Chipman, J.K. (2011) Comprehensive profiling of zebrafish hepatic proximal promoter CpG island methylation and its modification during chemical carcinogenesis. *BMC Genomics*, 12 (1): 1-16.
- Mischoulon, D., Fava, M. (2002) Role of *S*-adenosyl-L-methionine in the treatment of depression: A review of the evidence. *The American Journal of Clinical Nutrition*, 76 (5): 11585-11615.
- Mockler, T.C., Ecker, J.R. (2005) Application of DNA tiling arrays for whole-genome analysis. *Genomics*, 85 (1): 1-15.

- Moggs, J.G., Goodman, J.I., Trosko, J.E., Roberts, R.A. (2004) Epigenetics and cancer: Implications for drug discovery and safety assessment. *Toxicology and Applied Pharmacology*, 196 (3): 422-430.
- Mompalao, R.L., Bovenzi, V. (2000) DNA methylation and cancer. *Journal of Cellular Physiology*, 183 (2): 145-154.
- Moore, M.J., Smolowitz, R.M., Stegeman, J.J. (1997) Stages of hydropic vacuolation in the liver of winter flounder *Pleuronectes americanus* from a chemically contaminated site. *Disease of Aquatic Organisms*, 31: 19-28.
- Nakagawa, T., Yuan, J. (2000) Cross-talk between two cysteine protease families: Activation of caspase-12 by calpain in apoptosis. *The Journal of Cell Biology*, 150 (4): 887-894.
- Niculescu, M.D., Zeisel, S.H. (2002) Diet, methyl donors and DNA methylation: Interactions between dietary folate, methionine and choline. *The Journal of Nutrition*, 132 (8): 2333S-2335S.
- NMMP (national marine monitoring programme); Second report (2004) *Marine Environment Monitoring Group*, Cefas, UK.
- Norris, D.P., Patel, D., Kay, G.F., Penny, G.D., Brockdorff, N., Sheardown, S.A., Rastan, S. (1994) Evidence that random and imprinted Xist expression is controlled by preemptive methylation. *Cell*, 77 (1): 41-51.
- Oates, N.A., van Vliet, J., Duffy, D.L., Kroes, H.Y., Martin, N.G., Boomsma, D.I., Campbell, M., Coulthard, M.G., Whitelaw, E., Chong, S. (2006) Increased DNA methylation at the AXIN1 gene in a monozygotic twin from a pair discordant for a caudal duplication anomaly. *The American Journal of Human Genetics*, 79 (1): 155-162.
- Obrero, M., Yu, D.V., Shapiro, D.J. (2002) Estrogen receptor-dependent and estrogen receptor-independent pathways for Tamoxifen and 4-Hydroxytamoxifen-induced programmed cell death. *The Journal of Biological Chemistry*, 277 (47): 45695-45703.
- Okano, M., Xie, S., Li, E. (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics*, 19 (3): 219-220.
- Pardoll, R., Clarke, M.F., Morrison, S.J. (2003) Applying the principles of stem-cell biology to cancer. *Nature Reviews Cancer*, 3 (12): 895-902.
- Parkin, D.M., Bray, F., Ferlay, J., Pisani, P. (2002) Global cancer statistics, 2002. *A Cancer Journal for Clinicians*, 55 (2): 74-108.
- Paul, C.L., Clark, S.J. (1996) Cytosine methylation quantification by automated genomic sequencing and GENESCAN analysis. *Biotechniques*, 21 (1): 126-133.
- Pei, Y., Zhang, T., Renault, V., Zhang, X. (2009) An overview of hepatocellular carcinoma study by omics-based methods. *Acta Biochimica et Biophysica Sinica*, 41 (1): 1-15.

Peto, J. (2001) Cancer epidemiology in the last century and the next decade. *Nature*, 411 (6835): 390-395.

Phillips, J.M., Burgoon, L.D., Goodman, J.I. (2009b) Phenobarbital elicits unique, early changes in the expression of hepatic genes that affect critical pathways in tumour-prone B6C3F1 mice. *Toxicological Sciences*, 109 (2): 193-205.

Phillips, J.M., Goodman, J.I. (2009a) Multiple genes exhibit phenobarbital-induced constitutive active/androstane receptor-mediated DNA methylation changes during liver tumourigenesis and in liver tumours. *Toxicological Sciences*, 108 (2): 273-289.

Pogribny, I.P. (2010) Epigenetic events in tumourigenesis: Putting the pieces together. *Experimental Oncology*, 32 (3): 132-136.

Poirier, L.A., Wise, C.K., Delongchamp, R.R., Sinha, R. (2001) Blood determinations of *S*-adenosylmethionine, *S*-adenosylhomocysteine, and homocysteine: Correlations with diet. *Cancer Epidemiology, Biomarkers and Prevention*, 10 (6): 649-655.

Polakis, P. (2000) Wnt signalling and cancer. *Genes and Development*, 14 (15): 1837-1851.

Pomraning, K.R., Smith, K.M., Freitag, M. (2009) Genome-wide high throughput analysis of DNA methylation in eukaryotes. *Science*, 47 (3): 142-150.

Postlethwait, J.H., Yan, Y., Gates, M.A., Hornes, S., Amores, A., Brownlie, A., Donovan, A., Egan, E.S., Force, A., Gong, Z., Goutel, C., Fritz, A., Kelsh, R., Knapik, E., Liao, E., Paw, B., Ransom, D., Singer, A., Thomson, M., Abduljabbar, T.S., Yelick, P., Beier, D., Joly, J., Larhammar, D., Rosa, F., Westerfield, M., Zon, L.I., Johnson, S.L., Talbot, W.S. (1998) Vertebrate genome evolution and the zebrafish gene map. *Nature Genetics*, 18 (4): 345-349.

Prins, G.S. (2008) Estrogen imprinting: When your epigenetic memories come back to haunt you. *Endocrinology*, 149 (12): 5919-5921.

Probst, A.V., Dunleavy, E., Almouzni, G. (2009) Epigenetic inheritance during the cell cycle. *Nature Reviews Molecular Cell Biology*, 10 (3): 192-206.

Rajeevan, M.S., Vernon, S.D., Taysavang, N., Unger, E.R. (2001) Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *Journal of Molecular Diagnostics*, 3 (1): 26-31.

Ramakers, C., Ruijter, J.M., Lekanne, R.H., Antoon, F.M. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosciences Letter*, 339 (1): 62-66.

Ramsahoye, B.H. (2002) Measurement of genome wide DNA methylation by reversed-phase high-performance liquid chromatography. *Methods*, 27 (2): 156-161.

Rauch, T., Li, H., Wu, X., Pfeifer, G.P. (2006) MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent

methylation of homeodomain-containing genes in lung cancer cells. *Cancer Research*, 66 (16): 7939-7947.

Rauch, T., Wang, Z., Zhang, X., Zhong, X., Wu, X., Lau, S.K., Kernstine, K.H., Riggs, A.D., Pfeifer, G.P. (2007) Homeobox gene methylation in lung cancer studied by genome-wide analysis with microarray-based methylation CpG island recovery assay. *Proceedings of the National Academy of Science of the United States of America*, 104 (13): 5527-5532.

Raychaudhuri, S., Sutphin, P.D., Chang, J.T., Altman, R.B. (2001) Basic microarray analysis: Grouping and feature reduction. *Trends in Biotechnology*, 19 (5): 189-193.

Reamon-Buettner, S.M., Mutschler, V., Borlak, J. (2008) The next innovation cycle in toxicogenomics: Environmental epigenetics. *Mutation Research*, 659 (1-2): 158-165.

Reik, W., Walter, J. (2001) Evolution of imprinting mechanisms: The battle of the sexes begins in the zygote. *Nature Genetics*, 27 (3): 255-256.

Rein, T., DePamphilis, M.L., Zorbas, H. (1998) Identifying 5-methylcytosine and related modifications in DNA genomes. *Nucleic Acids Research*, 26 (10): 2255-2264.

Rhee, I., Bachman, K.E., Ho Park, B., Jair, K., Chiu Yen, R., Schuebel, K.E., Cui, H., Feinberg, A.P., Lengauer, C., Kinzler, K.W., Baylin, S.B., Vogelstein, B. (2002) DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature*, 416 (6880): 552-556.

Robertson, K.D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F.A., Jones, P.A. (1999) The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumours. *Nucleic Acids Research*, 27 (11): 2291-2298.

Rosen, N., She, Q.B. (2006) AKT and cancer--is it all mTOR? *Cancer Cell*, 10 (4): 254-256.

Rotchell, J.M., Lee, J., Chipman, J.K., Ostrander, G.K.. (2001a) Structure, expression and activation of fish ras genes. *Aquatic Toxicology*, 55 (1-2): 1-21.

Rotchell, J.M., Blairb, J.B., Shimc, J.K., Hawkinsd, W.E., Ostrandera, G.K (2001b) Cloning of the Retinoblastoma cDNA from the Japanese medaka (*Oryzias latipes*) and preliminary evidence of mutational alterations in chemically-induced retinoblastomas. *Gene*, 263 (1-2): 231-237.

Rozen, S., Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, 132: 365-386.

Safe, S., Kim, K. (2008) Nonclassical genomic ER/Sp and ER/AP-1 signalling pathways. *Journal of Molecular Endocrinology*, 41 (5): 263-275.

Sakinah, S.A., Handayani, S.T., Hawariah, L.P. (2007) Zerumbone induced apoptosis in liver cancer cells via modulation of Bax/Bcl-2 ratio. *Cancer Cell International*, 7(4): 1-11.

- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular cloning: A laboratory manual, Second edition. *Cold Spring Harbor Laboratory*, 1-1659.
- Sato, N., Maehara, N., Su, G.H., Goggins, M. (2003) Effects of 5-Aza-2-deoxycytidine on matrix metalloproteinase expression and pancreatic cancer cell invasiveness. *Journal of the National Cancer Institute*, 95 (4): 327-330.
- Schaefer, M., Pollex, T., Hanna, K., Tuorto, F., Meusburger, M., Helm, M. and Lyko, F. (2010) RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes and Development*, 24 (15): 1590-1595.
- Schmale, M.C., Gibbs, P.D.L., Campbell, C.E. (2002) A virus-like agent associated with neurofibromatosis in damselfish. *Disease of Aquatic Organisms*, 49 (107): 115.
- Schneider, S., Kaufmann, W., Buesen, R., Van Ravenzwaay, B. (2008) Vinclozolin—the lack of a transgenerational effect after oral maternal exposure during organogenesis. *Reproductive Toxicology*, 25 (3): 352-360.
- Schwab, M., Scholl, E. (1981) Neoplastic pigment cells induced by N-methyl-N-nitrosourea (MNU) in *Xiphophorus* and epigenetic control of their terminal differentiation. *Differentiation*, 19 (1-3): 77-83.
- Scott, A.P., Sanders, M., Stentiford, G.D., Reese, R.A., Katsiadaki, I. (2007) Evidence for estrogenic endocrine disruption in an offshore flatfish, the dab (*Limanda limanda*). *Marine Environmental Research*, 64 (2): 128-148.
- Sharma, S., Kelly, T.K., Jones, P.A. (2010) Epigenetics in cancer. *Carcinogenesis*, 31 (1): 27-36.
- Shheader, D.L., Williams, T.D., Lyons, B.P., Chipman, J.K. (2006) Oxidative stress response of European flounder (*Platichthys flesus*) to cadmium determined by a custom cDNA microarray. *Marine Environmental Research*, 62 (1): 33-44.
- Shi, H., Maier, S., Nimmrich, I., Yan, P.S., Caldwell, C.W., Olek, A., Huang, T.H (2003) Oligonucleotide-based microarray for DNA methylation analysis: Principles and applications. *Journal of Cellular Biochemistry*, 88 (1): 138-143.
- Shimoda, N. (2007) Progress in DNA methylation research: DNA methylation in zebrafish. *Nova Science Publishers, Inc*, 133-152.
- Shugart, L.R. (1990) 5-methyl deoxycytidine content of DNA from bluegill sunfish (*Lepomis macrochirus*) exposed to benzo[α]pyrene. *Environmental Toxicology and Chemistry*, 9 (2): 205-208.
- Sibani, S., Melnyk, S., Pogribny, I.P., Wang, W., Hiou-Tim, F., Deng, L., Trasler, J., James, S.J., Rozen, R. (2002) Studies of methionine cycle intermediates (SAM, SAH), DNA methylation and the impact of folate deficiency on tumour numbers in Min mice. *Carcinogenesis*, 23 (1): 61-65.

Siddiqi, M.A., Laessig, R.H., Reed, K.D. (2003) Polybrominated diphenyl ethers (PBDEs): New pollutants–old diseases. *Clinical Medicine and Research*, 1 (4): 281-290.

Siegfried, Z., Simon, I. (2010) DNA methylation and gene expression. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2 (3) 362-371.

Simpson, V.J., Johnson, T.E., Hammen, R.F. (1986) *Caenorhabditis elegans* DNA does not contain 5-methylcytosine at any time during development or aging. *Nucleic Acids Research*, 14 (16): 6711-6719.

Skinner, M.K., Manikkam, M., Guerrero-Bosagna, C. (2010) Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends in Endocrinology and Metabolism*, 21 (4): 214-222.

Small, H.J., Williams, T.D., Sturve, J., Chipman, J.K., Southam, A.D., Bean, T.P., Lyons, B.P. and Stentiford, G.D. (2010) Gene expression analyses of hepatocellular adenoma and hepatocellular carcinoma from the marine flatfish *Limanda limanda*. *Disease of Aquatic Organisms*, 88 (2): 127-141.

Sohn, B.H., Park, I.Y., Julee, J., Yang, S.-J., Jang, Y.J., Park, K.C., Kim, D.J., Lee, D.C., Sohn, H.A., Kim, T.W., Yoo, H.-S., Choi, J.Y., Bae, Y.S., Yeom, Y.I. (2010) Functional switching of TGF- β signalling in liver cancer via epigenetic modulation of a single CpG site in TTP promoter. *Gastroenterology*, 138 (5): 1898-1908.

Southam, A.D., Easton, J.M., Stentiford, G.D., Ludwig, C., Arvanitis, T.N., Viant, M. (2008) Metabolic changes in flatfish hepatic tumours revealed by NMR-based metabolomics and metabolic correlation networks. *Proteome Research*, 7 (12): 5277-5285.

Spitsbergen, J.M., Tsai, H.W., Reddy, A., Miller, T., Arbogast, D., Hendricks, J.D., Bailey, G.S. (2000) Neoplasia in zebrafish (*Danio rerio*) treated with 7,12-dimethylbenz[α]anthracene by two exposure routes at different developmental stages. *Toxicologic Pathology*, 28 (5): 705-715.

Stead, L.M., Au, K.P., Jacobs, R.L., Brosnan, M.E., Brosnan, J.T. (2001) Methylation demand and homocysteine metabolism: Effects of dietary provision of creatine and guanidinoacetate. *American Journal of Physiology-Endocrinology and Metabolism*, 281 (5): E1095-E1100.

Stehr, C.M., Johnson, L.L., Myers, M.S. (1998) Hydropic vacuolation in the liver of three species of fish from the U.S. West Coast: Lesion description and risk assessment associated with contaminant exposure. *Disease of Aquatic Organisms*, 32 (2): 119-135.

Stentiford, G.D., Bignell, J.P., Lyons, B.P., Feist, S.W. (2009) Site-specific disease profiles in fish and their use in environmental monitoring. *Marine Ecology Progress Series*, 381: 1-15.

Stentiford, G.D., Feist, S.W. (2005) First reported cases of intersex (ovotestis) in the flatfish species dab *Limanda limanda*: Dogger Bank, North Sea. *Marine Ecology Progress Series*, 301: 307-310.

Stouder C, Paoloni-Giacobino A. (2010) Transgenerational effects of the endocrine disruptor vinclozolin on the methylation pattern of imprinted genes in the mouse sperm. *Reproduction*, 139 (2): 373-379.

Strathdee, G., Brown, R. (2002) Aberrant DNA methylation in cancer: Potential clinical interventions. *Expert Reviews in Molecular Medicine*, 4 (4): 1-17.

Stromqvist, M., Tooke, N., Brunstrom, B. (2010) DNA methylation levels in the 5' flanking region of the vitellogenin I gene in liver and brain of adult zebrafish (*Danio rerio*)-sex and tissue differences and effects of 17 α -ethinylestradiol exposure. *Aquatic Toxicology*, 98 (3): 275-281.

Struys, E.A., Jansen, E.E., de Meer, K., Jakobs, C. (2000) Determination of S-adenosylmethionine and S-adenosylhomocysteine in plasma and cerebrospinal fluid by stable-isotope dilution tandem mass spectrometry. *Clinical chemistry*, 46 (10): 1650-1656.

Sumpter, J.P., Jobling, S. (1995) Vitellogenesis as a biomarker for estrogenic contamination of the aquatic environment. *Environmental Health Perspectives*, 103 (suppl 7): 173-178.

Sun, L., Hui, A., Kanai, Y., Sakamoto, M., Hirohashi, S. (1997) Increased DNA methyltransferase expression is associated with an early stage of human hepatocarcinogenesis. *Cancer Science*, 88 (12): 1165-1170.

Suzuki, M.M., Bird, A. (2008) DNA methylation landscapes: Provocative insights from epigenomics. *Nature*, 9 (6): 465-475.

Tahiliani, M., Peng Koh, K., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., Rao, A. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 324 (5929): 930-935.

Takai, D., Jones, P.A. (2002) Comprehensive analysis of CpG islands in human chromosome 21 and 22. *Proceedings of the National Academy of Science of the United States of America*, 99 (6): 3740-3745.

Tate, C.M., Lee, J., Skalnik, D.G. (2010) CXXC finger protein 1 restricts the Setd1A histone H3K4 methyltransferase complex to euchromatin. *The Federation of European Biochemical Societies Journal*, 277 (1): 210-223.

Taylor, K.H., Kramer, R.S., Davis, J.W., Guo, J., Duff, D.J., Xu, D., Caldwell, C.W., Shi, H. (2007) Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing. *Cancer Research*, 67 (18): 8511-8518.

Thomson, J.P., Skene, P.J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A.R.W., Deaton, A., Andrews, R., James, K.D., Turner, D.J., Illingworth, R., Bird, A. (2010) CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature*, 464 (7291): 1082-1087.

- Tilton, S.C., Givan, S.A., Pereira, C.B., Bailey, G.S., Williams, D.E. (2006) Toxicogenomics profiling of the hepatic tumour promoters indole-3-carbinol, 17 β -estradiol and β -naphthoflavone in rainbow trout. *Toxicological Sciences*, 90 (1): 61-72.
- Tischoff, I., Tannapfel, A. (2008) DNA methylation in hepatocellular carcinoma. *World Journal of Gastroenterology*, 14 (11): 1741-1748.
- Tsai, N., Wei, L. (2010) RhoA/Rock1 signalling regulates stress granules formation and apoptosis. *Cellular Signalling*, 22 (4): 668-675.
- Turner, B.M. (2009) Epigenetic responses to environmental change and their evolutionary implications. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 364 (1534): 3403-3418.
- Turner, B.M. (2007) Defining an epigenetic code. *Nature Cell Biology*, 9 (1): 2-6.
- UK marine monitoring and assessment strategy community (UKMMAS) (2010). Charting progress 2 feeder report: Clean and safe seas. (Eds. Law, R. and Maes, T.). *Department for Environment Food and Rural Affairs on behalf of UKMMAS*, 1-366.
- Ulrey, C.L., Liu, L., Andrews, L.G., Tollefsbol, T.O. (2005) The impact of metabolism on DNA methylation. *Human Molecular Genetics*, 14 (1): 139-147.
- Van Vlodrop, I.J.H., Niessen, H.E.C., Derks, S., Baldewijns, M.M.L.L., Van Criekinge, W., Herman, J.G., Van Engeland, M. (2011) Analysis of promoter CpG island hypermethylation in cancer: Location, location, location! *Clinical Cancer Research*, 17 (13): 4225-4231.
- Vandeghechuchte, M.B., De Coninck, D., Vandenbrouck, T., De Coen, W.M., Janssen, C.R. (2010a) Gene transcription profiles, global DNA methylation and potential transgenerational epigenetic effects related to Zn exposure history in *Daphnia magna*. *Environmental Pollution*, 158 (10): 3323-3329.
- Vandeghechuchte, M.B. and Janssen, C.R. (2011) Epigenetics and its implications for ecotoxicology. *Ecotoxicology*, 20 (3): 607-624.
- Vandeghechuchte, M.B., Kyndt, T., Vanholme, B., Haegeman, A., Gheysen, G., Janssen, C.R. (2009a) Occurance of DNA methylation in *Daphnia magna* and influence of multigeneration Cd exposure. *Environmental International*, 35 (4): 700-706.
- Vandeghechuchte, M.B., Lemièr, F., Janssen, C.R. (2009b) Quantitative DNA methylation in *Daphnia magna* and effects of multigeneration Zn exposure. *Comparative Biochemistry and Physiology*, 150 (3): 343-348.
- Vandeghechuchte, M.B., Lemièr, F., Vanhaecke, L., Vanden Berghe, W., Janssen, C.R. (2010b) Direct and transgenerational impact on *Daphnia magna* of chemicals with a known effect on DNA methylation. *Comparative Biochemistry and Physiology*, 151 (3): 278-285.
- Varriale, A., Bernardi, G. (2006b) DNA methylation in reptiles. *Gene*, 385: 122-127.

- Varriale, A., Bernardi, G. (2006a) DNA methylation and body temperature in fishes. *Gene*, 385: 111-121.
- Viant, M.R. (2007) Metabolomics of aquatic organisms: The new 'omics' on the block. *Marine Ecology Progress Series*, 332: 301-306.
- Vineis, P., Chuang, S., Vaissiere, T., Cuenin, C., Ricceri, F., The Genair-EPIC Investigators, Johansson, M., Ueland, P., Brennan, P., Herceg, Z. (2011) DNA methylation changes associated with cancer risk factors and blood levels of vitamin metabolites in a prospective study. *Epigenetics*, 6 (2): 1-7.
- Vogelestein, B., Kinzler, K.W. (2004) Cancer genes and the pathways they control. *Nature Medicine*, 10 (8): 789-799.
- Wade, P.A. (2001) Methyl CpG-binding proteins and transcriptional repression. *BioEssays*, 23 (12): 1131-1137.
- Wang, Y., Wang, C., Zhang, J., Chen, Y., Zuo, Z. (2009) DNA hypomethylation induced by tributyltin, triphenyltin, and a mixture of these in *sebastiscus marmoratus* liver. *Aquatic Toxicology*, 95 (2): 93-98.
- Ward, D.G., Wei, W., Cheng, Y., Billingham, L.J., Martin, A., Johnson, P.J., Lyons, B.P., Feist, S.W., Stentiford, G.D. (2006) Plasma proteome analysis reveals the geographical origin and liver tumour status of dab (*Limanda limanda*) from UK marine waters. *Environmental Science and Technology*, 14 (12): 4031-4036.
- Wardle, F.C., Odom, D.T., Bell, G.W., Yuan, B., Danford, T.W., Wiellete, E.L., Herbolsheimer, E., Sive, H.L., Young, R.A. and Smith, J.C. (2006) Zebrafish promoter microarray identify actively transcribed embryonic genes. *Genome Biology*, 7 (8): 1-14.
- Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., Schubeler, D (2005) Chromosome-wide and promoter-specific analyses identify sites of different DNA methylation in normal and transformed human cells. *Nature Genetics*, 37 (8): 853-862.
- Weinberg, R.A. (1989) Oncogenes, antioncogenes, and the molecular bases of multistep carcinogenesis. *Cancer Research*, 49 (14): 3713-3721.
- Weisenberger, D.J., Romano, L.J. (1999) Cytosine methylation in a CpG sequence leads to enhanced reactivity with benzopyrene diol epoxide that correlates with conformational change. *The Journal of Biological Chemistry*, 274 (34): 23948-23955.
- Whitelaw, N.C., Whitelaw, E. (2008) Transgenerational epigenetic inheritance in health and disease. *Current Opinion in Genetics and development*, 18 (3): 273-279.
- Whitelaw, N.C., Whitelaw, E. (2006) How lifetimes shape epigenotype within and across generations. *Human Molecular Genetics*, 15 (2): 131-137.

- Wild, C.P. (2009) Environmental exposure measurement in cancer epidemiology. *Mutagenesis*, 24 (2): 117-125.
- Williams, K., Christensen, J., Terndrup Pedersen, M., Johansen, J.V., Cloos, P.A.C., Rappsilber, J., Helin, K. (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature*, 473 (7347): 343-349.
- Williams, T.D., Diab, A.M., George, S.G., Godfrey, R.E., Sabine, V., Conesa, A., Minchin, S.D., Watts, P.C., Chipman, J.K. (2006) Development of the GENIPOL European flounder (*Platichthys flesus*) microarray and determination of temporal transcriptional responses to cadmium at low dose. *Environmental Science and Technology*, 40 (20): 6479-6488.
- Williams, T.D., Diab, A.M., George, S.G., Sabine, V., Chipman, J.C (2007) Gene expression responses of European flounder (*Platichthys flesus*) to 17-beta estradiol. *Toxicology Letters*, 168 (3): 236-248.
- Wilson, M.J., Shivapurkar, N., Poirier, L.A. (1984) Hypomethylation of hepatic nuclear DNA in rats fed with a carcinogenic methyl-deficient diet. *The Biochemical Journal*, 218 (3): 987-990.
- Wilson, V.L., Smith, R.A., Longoria, J., Liotta, M.A., Harper, C.M., Harris, C.C. (1987) Chemical carcinogen-induced decreases in genomic 5-methyldeoxycytidine content of normal human bronchial epithelial cells. *Proceedings of the National Academy of Science of the United States of America*, 84 (10): 3298-3301.
- World Cancer Research Fund (2007) Food, nutrition, physical activity and the prevention of cancer: A global perspective. *American Institute of Cancer Research, Washington*.
- Yan, Q., Huang, J., Fan, T., Zhu, H., Muegge, K. (2003) Lsh, a modulator of CpG methylation, is crucial for normal histone methylation. *The EMBO Journal*, 22 (19): 5154-5162.
- Yang, J., Liu, X., Bhalla, K., Kim, C.N., Ibrado, A.M., Cai, J., Peng, T.I., Jones, D.P., Wang, X. (1997) Prevention of apoptosis by Bcl-2: Release of cytochrome c from mitochondria blocked. *Science*, 275 (5303): 1129-1132.
- Yang, L., Li, N., Wang, C., Yu, Y., Yuan, L., Zhang, M., Cao, X. (2004) Cyclin L2, a novel RNA polymerase II-associated cyclin, is involved in pre-mRNA splicing and induces apoptosis of human hepatocellular carcinoma cells. *The Journal of Biological Chemistry*, 279 (12): 11639-11648.
- Yang, L., McBurney, D., Tang, S.C., Carlson, S.G., Horton, W.E. (2007) A novel role for bcl-2 associated-athanogene-1 (BAG-1) in regulation of the endoplasmic reticulum stress response in mammalian chondrocytes. *Journal of Cellular Biochemistry*, 102 (3): 786-800.
- Yeung, K.Y., Ruzzo, W.L. (2001) Principal component analysis for clustering gene expression data. *Bioinformatics*, 17 (9): 763-774.

- Yi, P., Melnyk, S., Pogribna, M., Pogribny, I.P., Hine, R.J., James, S.J. (2000) Increase in plasma homocysteine associated with parallel increases in plasma S-adenosylhomocysteine and lymphocyte DNA hypomethylation. *The Journal of Biological Chemistry*, 275 (38): 29318-29323.
- Yoder, J.A., Walsh, C.P., Bestor, T.H. (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics*, 13 (8): 335-340.
- Zhang, S., Wang, S., Li, H., Li, L. (2011) Vitellogenin, a multivalent sensor and an antimicrobial effector. *The International Journal of Biochemistry and Cell Biology*, 43 (3): 303-305.
- Zhou, X., Richon, V.M., Wang, A.H., Yang, X., Rifkind, R.A., Marks, P.A. (2000) Histone deacetylase 4 associates with extracellular signal-regulated kinases 1 and 2, and its cellular localization is regulated by oncogenic Ras. *Proceedings of the National Academy of Science of the United States of America*, 97 (26): 14329-14333.
- Zhou, X.W., Zhu, G.N., Jilisa, M., Sun, J.H. (2001) Influence of Cu, Zn, Pb, Cd and their heavy metal ion mixture on the DNA methylation level of the fish (*Carassius auratus*). *China Environmental Science*, 21 (6): 549-552.
- Zhu, H., Geiman, T.M., Xi, S., Jiang, Q., Schmidtman, A., Chen, T., Li, E., Muegge, K. (2006) Lsh is involved in *de novo* methylation of DNA. *The EMBO Journal*, 25 (2): 335-345.
- Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., Henikoff, S (2007) Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genetics*, 39 (1): 61-69.
- Zon, L.L. (2010) Zebrafish: A new model for human disease. *Genome Research*, 9 (2): 99-100.

Chapter 8

Appendix

8.1. List of publications

1. **Mirbahai, L.**, Williams, T.D., Chipman, J.K. (2012) Epigenetic memory: Its significance in regulation of responses to toxicants in aquatic species and potential as a marker of life-time exposures. Submitted to *Environmental Science & Technology*.
2. **Mirbahai, L.**, Yin, G., Bignell, J.P., Li, N., Williams, T.D., Chipman, J.K. (2011) DNA methylation in liver tumourigenesis in fish from the environment. *Epigenetics*, 6 (11): 1319-1333.
3. **Mirbahai, L.**, Williams, T.D., Zhan, H., Gong, Z., Chipman, J.K. (2011) Comprehensive profiling of zebrafish hepatic proximal promoter CpG island methylation and its modification during chemical carcinogenesis. *BMC Genomics*, 12: 1-16.
4. Chan, T., Williams, T.D., Mally, A., Hamberger, C., **Mirbahai, L.**, Hickling, K., Chipman, J.K. (2011) Gene expression and epigenetic changes by furan in rat liver. *Toxicology*. 292 (2-3): 63-70
5. **Mirbahai, L.**, Kershaw, R.M., Green, R.M., Hayden, R.E., Meldrum, R.A., Hodges, N.J. (2010) Use of a molecular beacon to track the activity of base excision repair protein OGG1 in live cells. *DNA repair*, 9 144-152

8.2. List of abstracts for oral presentations

1. **Mirbahai, L.**, Williams, T.D., Yin, G., Bignell, J.P., Li, N., Chipman, J.K. An epigenetic link between environment and cancer in fish liver tumours in the wild. Winter meeting of the UK molecular epidemiology group: Epigenetics and the environment. 14th December 2011, The University of Leeds, UK.
2. **Mirbahai, L.**, Williams, T.D., Chipman, J.K. DNA methylation profiling of fish tumours. Chromatin and chromosomes: The new frontiers of science. 29th June 2011, The University of Birmingham, UK.

3. **Mirbahai, L.**, Williams, T.D., Li, N., Bignell, J.P., Lyons, B.P., Stentiford, G.D., Chipman, J.K. Epigenetic change in liver tumours of the flatfish dab. 16th International symposium on pollutant response in marine organisms. 15-18 May 2011, California, USA.

4. **Mirbahai, L.**, Williams, T.D., Gong, Z., Lyons, B.P., Zhan, H., Bignell, J.P., Chipman, J.K. The first DNA methylation profiling of fish tumours. Biosciences Graduate Research School Symposium. 18-19th April 2011, The University of Birmingham, UK.

5. **Mirbahai, L.**, Williams, T.D., Gong, Z., Zhan, H., Chipman, J.K. Comparison of DNA methylation in carcinogenesis between fish and humans and effective strategies to investigate non-model organisms. Epigenetic and developmental programming conference. 21st-22nd March 2011, Newcastle, UK.

8.3. List of abstracts for poster presentations

1. **Mirbahai, L.**, Williams, T.D., Gong, Z., Zhan, H., Chipman, J.K. Basal and carcinoma associated DNA methylation profiles in zebrafish liver. 16th International symposium on pollutant response in marine organisms. 15-18 May 2011, California, USA.

2. **Mirbahai, L.**, Williams, T.D., Chipman, J.K. DNA methylation profiling of normal and liver tumours of aquatic species. British Science Week. 2010, Birmingham, UK.

3. **Mirbahai, L.**, Williams, T.D., Lyons, B.P., Southam, A., Stentiford, G., Small, H., Viant, M., Gong, Z., Chipman, J.K. Epigenetic change in fish carcinogenesis. Society of Environmental Toxicology and Biochemistry (SETAC) 20th Annual Meeting. 23-27 May 2010, Seville, Spain.

4. **Mirbahai, L.**, Aniagu, S.O., Williams, T.D., Chipman, J.K. Comparative quantification of genome-wide methylation in fish liver using reverse-phase high performance liquid chromatography (HPLC). 15th International symposium on pollutant response in marine organisms. 17-20 May 2009, Bordeaux, France.